

18 March 1982

Haig and Mr Malone's successor

The State Department should think hard before appointing a successor to Mr James Malone: first, it should redefine the job.

To be appointed Assistant Secretary at the State Department with responsibility for science and technology is to be offered a poisoned chalice. Few of the recent incumbents of the post have made a success of the job. Even strong-minded people such as Dr Dixie Lee Ray (in her time a doughty environmentalist and chairman of the Atomic Energy Commission as it used to be) have failed to make a success of it. Now Mr James Malone, the conservative lawyer appointed by President Reagan a year ago, has gone (see page 187) in a cloud of speculation that his error (of omission) is that he has not succeeded in delivering the new non-Carter policy on nuclear proliferation that the President has been promising since well before his election. Although it is easy to understand how a person whose links with the nuclear industry were widely advertised to Congress at the time of the confirmation of his nomination for the post should afterwards have failed to persuade congressmen of his impartiality on proliferation issues, the chances are that Mr Malone's failure is that of a man asked to tackle an impossible job. Before Mr Malone's successor is appointed, the State Department should give some thought to his job description.

That science and technology should be represented in some tangible form in the hierarchy of the State Department now goes without saying. International relations are now too important, and too intimately linked with science and technology, to be left exclusively to diplomats. Mr Malone's continuing responsibility for representing the United States at the Law of the Sea Conference is an example of what is expected of incumbents of this post. Carrying a new non-proliferation policy through the labyrinth of the government (Congress included) is another. Although problems of arms control are formally the responsibility of the Arms Control and Disarmament Agency, no assistant secretary worth his salt could fail to have opinions on these questions (and would be counted a failure if he did not, especially when as now the head of that agency must seem to many in the State Department to be Secretary of State in waiting). But there are also important and teasing questions to be decided about technological relations between the United States and other states. Just now, Japan has people at the State Department as exercised as the Soviet Union; beating back the tide of imported electronic goods on the American market is as preoccupying as deciding what should be done about the pipeline for natural gas between the Soviet Union and Western Europe or about scientists visiting the United States from supposedly unfriendly states.

The essential difficulty of the appointment is that it requires its incumbent at once to execute important parts of the administration's policy (negotiating an agreement on the Law of the Sea that will better suit the present administration's book, for example) and at the same time to provide advice on questions that are inherently subtle and perplexing. To make things worse, the job requires its holder to be able to win friends in Congress without being able to offer domestic patronage, and yet to be able to survive in the tricky waters of diplomacy (which means more than merely remembering to say "Thank you, Mr Chairman" after even an angry intervention in a United Nations committee meeting). The underlying dilemma, for Mr Alexander Haig as well as for Mr Malone's successor, is that the executive and advisory functions of the job are not easily compatible but that the conventions of Washington suppose that no mere

assistant secretary can expect his opinions to carry weight unless he can also demonstrate that he has executive responsibility for some part of the administration's programme.

There are two common solutions to such a problem, one of which is to appoint to such an impossible job a person so eminent that he or she can neglect or delegate administrative responsibilities. The other is to support somebody with no particular expertise by a high-level advisory committee which, however infrequently it meets, includes people who can be noisy on the telephone. The more imaginative but in the long run by far the most effective solution is to seek the best of both worlds — an incumbent whose reputation will ensure that his advice is listened to, and an advisory committee that will if necessary amplify what he has to say and some mechanism for making sure that the politics (domestic and international) are attended to. The State Department, having seen successive holders of this office sink beneath the waves, must now grit its teeth for the time when science and technology must be taken seriously — really seriously.

The case of non-proliferation is a good touchstone for the appointment of Mr Malone's successor. President Carter impaled himself unnecessarily on a policy about the proliferation of nuclear weapons that seemed to non-nuclear powers a splendid illustration of the cryptic imperialism of the United States and to would-be American suppliers of nuclear equipment to be a restraint on their legitimate trade. The Reagan Administration is right to look for a better balance, one that will acknowledge that it is no more possible to outlaw the proliferation of nuclear weapons by an act of Congress than it is to repeal the second law of thermodynamics by executive order. Whether it will find such a solution will depend on its success in persuading Congress that looking for loopholes in the International Atomic Energy Agency's safeguards system is less important than finding ways of dissuading potential nuclear powers from making bombs. The trouble is that doing that is as much dependent on the old-fashioned skills of diplomacy as on what new technology has to offer. In short, the State Department should resolve to be less formal in its requirements of a policy on non-proliferation; it should appoint a successor to Mr Malone who will help it to be just that.

Uneconomic savings

The financial squeeze on British universities is to continue. Structural change is the only defence.

The British Government is nothing if not resolute in its determination to spend less on higher education. The annual forecast of public expenditure which accompanied last week's budget shows that total expenditure will continue to decline even after the academic year 1983-84 — that by which the 8.5 per cent reduction decreed at the end of 1980 will have been completed. Because all sectors of higher education are lumped together in the figures, it is not possible to tell how the continuing contraction will be shared between the universities and other kinds of institutions, the polytechnics for example. But one thing is certain — if the government's plans stick, student numbers will fall further until the proportion of young people in higher education is little more than 11 per cent. Especially when the demand for



higher education is increasing as quickly as at present, and when the government has itself provided for an increase in the cost of what is called "non-advanced further education" (vocational courses of various kinds), it makes no sense at all that, for the first time in British history, access to universities and to higher education in general should be restricted by decree of the Department of Education and Science.

Not even the British Government can be content with this illiberal state of affairs. Going down in history (and into a general election, now not more than two years away) with such a reputation cannot be much of a comfort. But those who will be permanently damaged are not the politicians responsible for these developments but educational institutions, universities and polytechnics, and the students whom they are now prevented from teaching. This is why it is simply not good enough that those who might be thought to be responsible for the long-term interests of the universities should for practical purposes be sitting on their hands, doing nothing to bring about the structural changes in higher education on which survival must ultimately depend. The University Grants Committee, which is technically responsible for foisting student quotas on its dependants, shows very little sign of advocating the changes in the mechanisms of financial support for universities that will be necessary if flexibility is ultimately to be achieved. The Committee of Vice-Chancellors and Principals, united (and successful) in its efforts in the past six months to establish that only the government can meet the cost of firing academics from tenured posts, is constitutionally incapable of pressing for arrangements that would encourage diversity within the university system. And the polytechnics, the other important sector of British higher education, are uncomfortably squeezed financially between local and central government yet as firmly isolated as ever from the rest of higher education. They suffer from the awkward circumstance that their bills are paid by local authorities, which are then promptly reimbursed by the central government. After two years of argument, the central government has wrested back only a token of control.

The folly of these arrangements is already clear. The student quotas to which universities are being required to plan for the academic year beginning eighteen months from now will ensure that almost all universities become smaller and thus make less efficient use of their capital equipment. Some will become so much smaller that they will be forced to abandon teaching in many fields and thus cease to be universities except in name. The obvious solution (see *Nature* 18 February, p.541) is to devise a mechanism that will enable the total cost to the government of supporting universities and their students to be defined in advance, and limited, without restricting the capacity of universities to perform the social functions of educating young people for which they were originally created. Such a device would also provide universities with an incentive towards diversity; universities fearful of their capacity to compete for students with, say, Oxbridge would be faced with the prospect either of making themselves distinctive or going out of business. Universities least able to compete would also find themselves persuaded by objective pressures to teach more economically or even to form associations among themselves whose effect would be an economical division of labour. What would be wrong with such a state of affairs except the unwillingness of individual universities to acknowledge that while they are titularly equal, some are more equal than others?

The practical obstacle to such an arrangement, however, is not so much the unwillingness of universities to acknowledge what must be done as the division between the two halves of the binary system — the universities on the one hand and the polytechnics created as a kind of academic counterforce by the late Mr Anthony Crosland in 1966. Both halves are now under severe pressure but independently. When academics are fired or teaching departments closed in either half of the binary system, it is in principle possible to ensure that the weakest will be the first to go. But there is no conceivable means by which the performance of people or departments in one half of the system can be compared with that of their equivalents in the other half. As a result, there is

every likelihood that the decisions now being made by universities and polytechnics separately are inconsistent, even nonsensical. Few among British taxpayers (many of whom are academics) dispute the need that the government should contain its public expenditure, but it is just as important that the cash limits now being enforced for higher education should be accomplished in the most economical way — that which will yield the greatest benefit for whatever funds there are to spend. After a long period during which the government has had to fight the local authorities for a measure of central control of the polytechnics, it is understandable that it should lack the stomach for trying to set up some common framework within which both kinds of institutions could function. But nothing less than that will meet the present urgent need.

Turkish illiberality

The scientific community must make up its mind how to help members denied freedom.

The sad case of Dr Yeter Ögelman (see *Nature* 25 February, p.638 and this issue, p.231), the Turkish scientist who spent more than half of last year in a Turkish gaol, is still unresolved. At the resumed hearing of her case last week, Dr Ögelman was neither acquitted nor found guilty of the charge that her part in the organization of a women's rights campaign in the early 1970s was subversive of the new social order in Turkey. The result is that the charge will remain in abeyance, allowing the authorities to breathe new life into it whenever they choose. Meanwhile, Dr Ögelman remains dismissed from her post in the physics department of Cukurova University.

What, in these circumstances, should the scientific community do? The first temptation is to protest that it is thoroughly illiberal, and a serious infringement of what has come to be considered ordinary people's rights, that a person should be hounded and imprisoned as Dr Ögelman has been. The difficulty, unfortunately, is that the scientific community is not competent to raise such a complaint. For Dr Ögelman has been prosecuted not because she is a physicist but because of her part in an organization intended to secure rights for women in a society in which these have recently been denied. So there is no cause for the scientific community or its academies to stand up and complain at Dr Ögelman's treatment. Exactly the same principle applies to other states, of which there are now all too many, in which scientists and others must work under conditions in which personal liberty is unreasonably restricted. Other people's laws may be distasteful, but they are, after all, laws from which people cannot be exempted because they are scientists.

So is there nothing that can be done for Dr Ögelman, and for the growing band of scientists imposed upon by martial law in Poland (and for that matter Argentina)? Formal protests are inappropriate — which is not to say that individuals should not be free to support other organizations that are competent to complain — but practical help is both appropriate and likely to be welcome. The obvious and shining example of what can be accomplished is that of the 1930s, when many of those for whom life had become intolerable in central Europe were able to find jobs and continue their careers in the West. Now, the plight of those whose freedom is constrained is rarely as desperate as in the years before the Second World War, and there is also something in the view that the cause of personal liberty itself will be helped if those in trouble with the authorities stay in their own countries. But the provision of refuges overseas is not the only practical help that can be offered and that is likely to be welcome. People who have lost their jobs can be helped to remain active in scientific work even from a distance. Grant-making bodies and those who organize international conferences can usefully consider the interests of their colleagues who are deprived of freedom. Is it too much to ask that the scientific community, rightly standing aloof from the internal affairs of states elsewhere, should nevertheless more systematically consider how to help those of its members who are in trouble?

Overseas science shake-up in US

Malone goes, Law of Sea again in limbo

Washington

The US State Department is carrying out a shake-up of some of its top officials responsible for international scientific and technological affairs. Last week, the department confirmed that Mr James L. Buckley, a former Republican senator who is now Under-Secretary of State for Security Assistance, Science and Technology, will become a direct assistant to Secretary of State Alexander Haig.

Mr Buckley's move coincides with the news, also confirmed by the department, that Mr James Malone is being removed from his post as Assistant Secretary of State for Oceans and International Environmental and Scientific Affairs, although he will remain head of the US delegation to the Law of the Sea negotiations which reopened in New York last week.

No official reasons have been given for the reshuffle. In the case of Mr Malone, who was previously a lawyer in private practice and had been general counsel of the Arms Control and Disarmament Agency under the Nixon Administration, sources in Washington suggest that there had recently been criticism of some of his staff appointments. It is also felt that Mr Malone had lacked sufficient time to provide adequate attention to a range of controversial policy areas for which his post is responsible.

The moves are also part of a general reorganization of the State Department which has followed the promotion of Deputy Secretary of State William Clark to the President's National Security Advisor. One result is expected to be a more pragmatic and less ideological approach to decisions involving technology and foreign policy.

The contraction of Mr Malone's responsibilities seems, in part, to be related to his failure to secure congressional backing for the proposed relaxation of restrictions on the export of nuclear technology.

Mr Malone, whose nomination to the State Department was backed by members of the right wing of the Republican party, and who has contributed to reports published by the conservative Heritage Foundation, ran into strong opposition during his nomination hearings over his previous links with the nuclear industry.

At the time it was revealed that, as a private lawyer, he had represented the nuclear industries of both Taiwan and Japan in negotiations with the United

States over access to nuclear technology. He promised to distance himself from any policy decisions involving former clients, but his links with the industry raised questions about how far his efforts to reformulate the US position on nuclear non-proliferation reflected the views of the industry rather than the foreign policy interests of the nation as a whole.

Whether or not the charge was a fair one, it has limited the effectiveness with which Mr Malone has been able to argue the Administration's position. "Once his nomination had been approved, he did not help matters by producing position papers that clearly looked like industry documents," an aide to Senator John Glenn, one of the sharpest critics of Mr Reagan's efforts to boost the exports of the US nuclear industry, said last week.

Although no announcement has been

made about Mr Malone's successor, it is widely expected that several of his responsibilities for nuclear export policy will be transferred to Mr Richard Kennedy, Under-Secretary of State for Management as well as ambassador to the International Atomic Energy Authority.

Mr Kennedy was a member of the Nuclear Regulatory Commission between 1976 and 1980. Before that he served for five years on the staff of the National Security Council where he worked directly under Dr Henry Kissinger, and has therefore had wide experience of both nuclear and foreign policy.

Mr Malone's reassignment could also have an impact on the outcome of the Law of the Sea negotiations, where the United States is at the centre of two separate conflicts, on the one hand expressing its strong opposition to some of the provisions in the

European fusion sustained

Brussels

Funds to be devoted to Europe's research effort in controlled thermonuclear fusion were increased from 900 to about 1,400 million European Currency Units (1 ECU = £0.56) at a meeting of the EEC's science ministers in Brussels last week. JET, the Joint European Torus, which is being built at Culham in the United Kingdom will have a 319 million ECU budget for 1982-86 when the construction phase will be completed. European Community finance for other European fusion research has been set at 301 million ECU. This, with the contributions from national research budgets, will bring European fusion spending up to 1,400 million ECU over the next five years. The contribution from the EEC's budget for JET alone is thus increased from 195 to 319 million ECU.

In fact, most details of the fusion programme had been worked out well before the ministerial council; however, the British did their best to cut down support for general fusion research to 280 million ECU, a long way from the 325 million ECU originally requested by the European Commission, by arguing that the Commission had overestimated the speed of new developments. The European Commission calculated that 80 scientists would be needed to work on planning NET, the Next European Torus, and that accounted for much of the increased budget. But UK officials thought that because countries such as the United States or Japan were making do with 50 planners, the EEC could also do with fewer. The 301 million ECU finally agreed are supplemented by national contributions for projects involving tokamaks smaller than JET and other types of plasma confinement systems. The timetable for JET

foresees its tokamak in operation by the end of 1986.

At the press conference held by Philippe Maystadt, the Belgian science minister, who chaired the council, speculation was aroused by the omission of further international cooperation with the Soviet Union. Cooperation with Japan and the United States is to be strengthened during the new programme, he said. All four fusion "powers" are, however, cooperating with the International Atomic Energy Agency in Vienna on the Intor project, a post-JET tokamak, originally proposed by the Soviet Union. Cooperation among the Western powers is taking place simultaneously within the International Energy Agency in Paris and this may now be strengthened. The European Commission refutes the idea that an iron curtain is to be drawn across fusion research as a result of the present political tensions, since it is a field which has traditionally been kept outside East-West politics.

The meeting also approved a research programme on raw materials for 54 million ECU which covers a wide variety of different research projects including research in metals and minerals, wood, the recycling of non-ferrous metals and ceramics.

The rest of the meeting was devoted to discussing two policy papers on the re-organisation of European research cooperation resulting from the meeting held last November. Among the ideas now being given serious consideration is the setting up of a Perception and Assessment Committee, which would bring together about twenty senior members to analyse the trends of science technology and guide EEC projects to counterbalance any deficiencies.

Jasper Becker

current draft of a proposed international treaty, on the other not wishing to provoke a split with other industrialized nations which, although expressing reservations, are less inclined to reject the treaty as it now stands.

Last Thursday, the US delegation presented the conference with a list of changes that it says must be made before it will sign the treaty. These include the abandonment of efforts to place limits on the amounts of different minerals that can be retrieved from the sea bed, a revised voting procedure that would give the United States greater influence over the actions of an international authority established to regulate deep-sea mining, and dropping the requirement that US mining companies should provide Third World consortia with the latest mining equipment.

At the beginning of what is expected to be the final round of negotiations started in the early 1970s, several non-aligned countries indicated that it was too late to re-open discussion on topics which, they claimed, had been settled at earlier negotiating sessions. These countries said they would proceed with drawing up a treaty, leaving the United States to decide whether or not it wants to sign.

Behind the rhetoric, however, the real question is how far each side is prepared to compromise. A treaty without a US signature, or US financial support for the regulatory institutions which would be set up, would be of limited impact. From the American point of view, some prospective mining companies are concerned that they might be unable to raise credit from multinational banks if the treaty went ahead without a US signature, while several industrialized countries — including France, Britain and West Germany — are reluctant to jeopardize their trading relationships with Third World nations by providing too much support for the United States.

David Dickson

Polish higher education

Reform reversed

Research projects in Poland are now to be "rechecked" to assess their usefulness to the national economy, according to Dr Benon Miskiewicz, the new Minister of Science, Higher Education and Technology. Furthermore, the system for appointing assistant lecturers will be modified, and new principles worked out for the employment of graduates.

Dr Miskiewicz was addressing a meeting of lecturers from higher educational institutes subordinate to his ministry. The gathering did not include medical schools or specialist bodies such as the Mining and Metallurgy Institute of Krakow which, last September, had been a principal focus of resistance to attempts by the then minister, Mr Jerzy Nawrocki, to renege on the academic autonomy which, it had been

promised, would be incorporated into the new higher education bill. This bill, much delayed by the imposition of martial law, is now apparently undergoing redrafting and modification before going forward to the Sejm (parliament).

Dr Miskiewicz's emphasis on the "usefulness" of research suggests a reemphasis of the "problem-orientated" research structure introduced in 1973 by which all research had to be associated with a specific problem of the national economy. This led to an unseemly scramble by researchers to find a "problem" rated as highly as possible by the planners. During the open press debates last summer on the future of Polish science, it was generally admitted that the "problem" structure had failed and should be replaced by some type of direct contact with industry.

By tacitly repudiating such proposals, Dr Miskiewicz was apparently trying to dissociate his ministry from such proposed reforms. During the past three months, there has been some high-level disagreement as to who was primarily responsible for the outspokenness of the academic community — the Academy of Sciences or the universities. Talking to the editorial board of the weekly magazine *Polityka* last January, Deputy Premier Mieczyslaw Rakowski said that the breaking point had come when the Conference of University Rectors had demanded a voice in the appointment of the Minister of Science, Higher Education and Technology. The conference, established in September 1980, was always regarded with suspicion by the government, and shortly after the declaration of martial law it was dissolved — unlike the majority of organizations which were simply suspended.

Officials of the Ministry of Science, Higher Education and Technology, however, blame the Academy of Sciences, whose institutes, they say, harboured dissidents who had been banned from university teaching. The proposed new law on the academy, which would have made the academic secretary responsible only to his fellow members and not, as at present, to the prime minister, has been abandoned. The limited autonomy which the academy has until exercised over its research budget, is likely to be reduced even further.

The authorities, however, seem unwilling to risk any further direct confrontation with the academy. A recent meeting of the presidium of the academy was addressed by Dr Hieronim Kubiak, the Communist Party spokesman on higher education, who is reputed to be the most liberal member of the Politburo. Dr Kubiak warned the academicians against pointless confrontation, stressing that the intelligentsia should "advocate a Poland which is possible". Nevertheless, he said, if there are reforms which are actually possible at present, one should support them openly and actively. **Vera Rich**

Industrial lure

New York

As five university presidents prepare to discuss the growing connection between universities and commercial interests in recombinant DNA research (*Nature* 4 March, p.1), private biotechnology firms continue to have an increasing impact. This week, two top academics decided to give up basic research posts to join private companies.

Dr Walter Gilbert, 1980 Nobel prize winner in chemistry, is to leave his American Cancer Society professorship at Harvard University as of June to retain his position as executive officer of Biogen, the Geneva- and now Cambridge-based biotechnology firm.

Dr Gilbert has had a two-year leave of absence from Harvard to fulfil his position as chairman of the board of directors of Biogen. His hopes of permanently maintaining both his laboratory at Harvard and his position at Biogen ran into unresolvable problems with the Harvard administration. He will, however, remain a senior associate of the university for two more years. His Harvard laboratory will continue to operate during this time.

His full-time commitment to Biogen comes at a time when the company's new American laboratory in Cambridge, Massachusetts has recently adopted another academic, Dr Richard Flavell from the National Institute of Medical Research at Mill Hill, as research director. Dr Flavell is officially on two-year leave of absence from Mill Hill.

Another top American academic to move into the private sector, Dr Edward Scolnick, chief of the laboratory of tumour virus genetics at the National Cancer Institute, has decided to become Merck, Sharpe and Dohme's executive director of virus and cell biology.

Michael Stein

UK budget

A gentle push

The British government's enthusiasm for new technologies, and its lack of enthusiasms for the universities, was confirmed in last week's budget. While the universities can expect little respite from the austerity promised a year ago, the development of new technologies in industry is to receive a modest boost. British Telecom, which is to be allowed to increase its annual capital expenditure by a half over the next three years, is the chief beneficiary. But the Chancellor of the Exchequer has also promised the Department of Industry an extra £130 million for new technology in the next financial year, 1982-83.

British Telecom, which has recently improved its profits, will now be able to invest £2,300 million on capital projects in 1982-83, rising to nearly £3,000 million in 1984-85 mostly from its own resources.

British Telecom is now adapting its business to the liberalization announced by the government last October, the chief effect of which is to allow other companies to attach equipment to the telephone network. The government is, however, now apparently considering taking the liberalization a step further by selling off part of British Telecom. The Department of Industry would not comment last week on suggestions that a new telecommunications bill is being drafted for debate in the autumn. But ministers are discussing a document outlining several options for future telecommunications policy.

The Department of Industry is also in the throes of deciding how to spend its own windfall in last week's budget. Around £30 million of the £130 million promised for new technology projects this year will be spent on teletex, a superior type of telex. Competitors for the remaining £100 million include programmes to support the use of microprocessors in industry and research on flexible manufacturing systems, information technology and electronics.

The department's space budget, however, is likely to receive at least some of the extra money, possibly £5 million a year over the next three years. **Judy Redfearn**

Gene manipulation watchdog

Too smooth?

The uniquely British way of regulating research in genetic manipulation, the Genetic Manipulation Advisory Group (GMAG), may yet be a victim of its own success. The immediate difficulty is that its work seems now to have become so uncontroversial that some of its sponsors and even some of its members are asking whether it should continue to exist. A decision one way or the other is likely by the end of the year, when the chairman of the group, Sir Robert Williams, and several of its members come to the end of their term of office.

Since GMAG three years ago adopted a method of qualitative risk assessment for categorizing experiments in genetic manipulation, at least its academic customers appear to have been satisfied with the informality of the procedures. Most academic work in the field appears to require only that laboratories should follow good microbiological practice or, alternatively, the lowest category of safety containment.

While notification of experiments is still legally required, this may be retrospective (in an annual report) for work carried out in open laboratories, while advance approval is no longer required for the

majority of experiments in the lowest category of containment. Notifications of experiments in the second level of containment appear to amount to only about half a dozen a year.

For practical purposes, the burden of assessing the safety of experiments has been devolved from GMAG to laboratory safety committees. The bulk of GMAG's work seems now to consist of the assessment of safety arrangements in laboratories newly embarking on genetic manipulation and the drafting of guidance notes — a revised (and relaxed) procedure for dealing with manufacturing processes appeared only last week.

One substantial problem remaining is the systematic categorization of experiments involving viruses which are not covered by the regulations on the laboratory use of dangerous pathogens but which would involve hazards not now foreseen — oncogenic viruses for example. GMAG is wondering whether to set up a committee to look into this question. But the group has shied away from the field of human embryo transplants on the grounds that such work does not fall within the statutory definition of genetic manipulation and because its members do not have the necessary expertise.

Although both the Confederation of British Industry and the Trades Union Congress, which nominate members of the



group, still mutter that regulation of genetic manipulation might advantageously be transferred to the Health and Safety Executive, which has a statutory responsibility for safety at work, this view is no longer being strongly promoted.

Academic clients of GMAG, on the other hand, would prefer that responsibility should remain with the group, which has become the devil that they know. Academic investigators would also prefer that GMAG should continue in being so as to deal with whatever unforeseen problems may arise.

The most serious threat to the continued existence of GMAG seems to be *ennui*, the lack of obvious challenge. Keeping the group alive is a modest public expense, making it a target for would-be economizers, but abolition would presumably require legislation — and parliamentary time is as scarce as money.

US electronics industry

Pooling effort

Washington

Faced with increasingly intense competition from the Japanese electronics industry, a number of large American computer companies are developing plans to pool a proportion of their research resources into a new venture, provisionally called Microelectronics and Computer Technology Enterprises (MCTE).

The idea is to achieve economies of scale for the companies, many of which invest up to 10 per cent of their annual revenues in research, by jointly funding research projects in such areas as materials science, computer-aided design, manufacturing processing and quality control.

Sixteen companies met in Florida last month to discuss plans for such a venture, which would have a research budget expected to rise eventually to about \$100 million a year. The meeting was also attended by representatives of the Department of Defense and the

UN assists India

New Delhi

The scheme developed by the United Nations Development Programme (UNDP) to encourage Indian scientists settled abroad to return to temporary work in India seems to be making progress. The TOKTEN scheme (Transfer of Know-how Through Expatriate Nationals) got under way early last year (see *Nature* 29 January 1981, p.343) and since then 32 Indian scientists from the United States, West Germany and Canada have visited India to work with industrial units and research institutions.

India has recently become alarmed at the extent of its 'brain drain', especially among scientific personnel. Although there are no official figures, estimates place the number settled abroad at more than 100,000.

The TOKTEN scheme, already hailed as a success in Turkey, was introduced in India, with an initial contribution of \$1.1 million to encourage scientists settled abroad to give the benefit of their knowledge to their homeland. UNDP pays the fares of those who agree to visit India for a period of two or three weeks to help solve technical problems in sectors chosen by the government. The project, which is being implemented by the Council of Scientific and Industrial Research, has evoked a favourable response from both the Indian experts and the organizations with which they have worked, many of which have asked that these scientists should make a return visit or that more should come forward.

Sunil Saraf

Massachusetts Institute of Technology, both of which have indicated their possible interest in collaborating with the joint research effort.

Pressure to find ways of linking the generic research needs of different companies in the computer industry has increased as the costs of research — both in real terms and as a proportion of revenues — has continued to rise.

A report published in 1979 by a panel of computer scientists established by the National Science Foundation suggested that a foundation should be established to raise funds from electronics companies and distribute them to university research departments for carrying out basic research.

As a step in this direction, the California-based Semiconductor Industries Association has established a Semiconductor Research Cooperative (ISRC) to act as a clearing house for companies interested in sponsoring university research. The idea of a separate foundation, however, has remained on the shelf.

Mr William Shaffer, spokesman for the Control Data Corporation which organized the Florida meeting, said last week that the ISRC initiative did not go far enough for the industry and was more limited in scope than the proposal which was discussed in Florida which would embrace not only a range of research activities designed to enhance the production of electronics and semiconductor components — an area in which the United States has lost considerable ground to Japan in recent years — but also provide training and other kinds of assistance to collaborating companies.

Computer manufacturers present at the meeting included Burroughs Corporation, Digital Equipment Corporation, Honeywell Incorporated, Sperry Corporation and Rand Corporation. Among the semiconductor producers were Advanced Micro Devices Incorporated. IBM had been invited to participate but did not attend.

The companies agreed to a general framework for a joint research venture that would embrace both basic and applied research. Such a joint venture will have to avoid conflict with strict American anti-trust laws, which have previously been quoted as a major impediment to collaborative research between different companies in the same industry. The need to "clarify" the application of anti-trust laws to research was one conclusion of an eighteen-month review of domestic policies affecting technological innovation carried out under President Carter; a subsequent report from the Department of Justice, published in November 1980, concluded that "in general, the closer the joint activity is to the basic research end of the research spectrum . . . the more likely it is to be acceptable under the anti-trust laws".

Control Data Corporation says that the outline of its proposals already presented

to the Federal Trade Commission has not met with any objections. The company is hoping that, by stressing the intention of MCTE to carry out research rather than produce electronics components, any potential conflict with the anti-trust laws will be minimized.

David Dickson

UK-Romania cooperation

One-way street

Bucharest

The meeting earlier this month of the UK-Romanian Joint Commission on Trade and Technology produced a novel suggestion from the Romanian delegation: the establishment of long-term joint technological research projects. But the suggestion is unlikely to stimulate a positive response from the British side, since attempts at cooperation over the past decade have almost all become bogged down in bureaucracy and delay.

Even the one moderately viable project — the production under licence in Romania of BAC-111 aircraft — has run into difficulties; and hopes that the finished product might be exhibited at this year's Farnborough air show are rapidly fading.

Joint research and development ventures are impeded not simply by the currency problems endemic in all dealings with Eastern Europe but also by the reorganization of Romanian science during the seventies which has been specifically aimed at eliminating dependence on foreign licences and the development of a native technological base. According to Romanian planners, during the past five years, Romanian research has provided some 90 per cent of the new materials and more than 90 per cent of the new technologies put into production. This appears to leave little room for technological flow and exchange.

Romanian-British cooperation is formally of two kinds. Pure research, including exchange scholarships, is handled by an agreement between the Royal Society and the Academy of Romania. The restructuring of Romanian science has, however, reduced the academy's role.

The driving force in Romanian science is now the National Council for Science and Technology, headed by Dr Elena Ceaucescu, wife of the president, and the main research potential is concentrated in ten "Central Research Institutes" whose emphasis is largely on applications. At the height of the reforms, in the mid-1970s, it seemed that fundamental research might vanish altogether in Romania; now, however, 10 per cent of the science budget is allotted to "fundamental research".

Although at the academic level it has been possible over the past few years to arrange some cooperation projects — such as the exchange agreement between the Polytechnic University of Bucharest and

the Polytechnic of Central London, the types of programmes that would normally fall within the competence of the joint commission are considerably harder to arrange. A joint seminar, held in Bucharest last May, on problems of energy saving in agriculture was — from the Romanian point of view — highly successful, but the British side is reluctant about a return visit, since they felt the "pooling" of knowledge came almost entirely from the British side.

Romanians involved in such ventures admit the disparity. But such one-side arrangements can never prove satisfactory in the long term to the "donor" partner, particularly when all procedural matters are hamstrung by bureaucracy. For, according to a law of 1974, any Romanian scientist contacted by a foreigner must report the matter to the relevant ministry. Although a number of prominent scientists say this regulation is a dead letter, the atmosphere of delay and obfuscation is hardly conducive to international cooperation.

Vera Rich

Bio-fund open

Those wishing to take advantage of the European Community's offer to support research and training in bio-molecular engineering should think fast. Applications for research contracts are needed by 15 May, and for training contracts by July.

After a long and arduous gestation period, the EEC sponsorship of bio-molecular engineering was looked on as something of a disappointment — "we came out with a mouse", said Vicomte Etienne Davignon, commissioner for research and development in Brussels (*Nature* 28 January, p.273).

But for all that, there is about 8 million European Currency Units (£4.5 million) available, and in the United Kingdom the Department of Industry shows signs of being anxious to alert UK scientists to the possibilities of obtaining EEC support for their work.

Research contracts are available for projects mainly related to agriculture and the safety of genetically manipulated organisms. Training contracts lasting 1 to 2 years from January 1983 cover a wider range of subjects including the development of cloning systems and gene transfer in agriculturally important plants. The commission invites submissions from "employed scientists (including those employed by industries) or unemployed scientists (who recently completed a doctoral thesis or postdoctoral assignment) who have demonstrated their capacity for high level scientific research". Applicants should approach the European Commission, but the Department of Industry will advise UK scientists how to apply.

Charles Wenz

Congress tackles 1983 budget

Close vote on science budget a surprise

Washington

In the first congressional action on the research and development budget for the fiscal year 1983 submitted by President Reagan at the beginning of February, a subcommittee of the House of Representatives Science and Technology Committee agreed by one vote last Thursday to recommend that the budget for the National Science Foundation (NSF) be increased by \$30 million over the President's request to a total of \$1,099.5 million.

The subcommittee left untouched the proposal to spend \$1,052.3 million on "research and related activities" sponsored by NSF, although it has suggested some reshuffling between categories. For example, it suggests cutting from 9.5 to 5.6 per cent the proposed increase in funding for mathematical and physical sciences, and that biological, behavioural and social sciences should grow by 14.2 rather than 6 per cent, particularly in the light of the major cuts in social science research last year.

The subcommittee has also proposed transferring some of the funds for research to support the International Institute for Applied Systems Analysis in Vienna. However, the Reagan Administration's decision to eliminate its contribution to the institute, originally taken last year but since reconsidered, is said to have been made final at a meeting of the National Security Council last week.

The main difference between the subcommittee's and the Administration's proposals are over NSF support for science education. Whereas the Reagan Administration, committed to phasing out such activities within NSF, has proposed a budget of \$15.0 million in 1983 — down from \$70.7 million in 1981 — the members of the subcommittee have registered their opposition by adding another \$30 million.

It was primarily this decision that led the Republican members of the subcommittee, as well as newly-elected Democrat Dave McCurdy of Oklahoma, to vote against the proposed increases. Disagreement over the science funds is also expected to generate another close vote in the full science committee when the budget authorization bill comes up for approval.

The closeness of the vote, especially among the members of what is generally considered to be one of the most pro-science subcommittees in Congress, suggests that the final figures for government support of research and development

next year are likely to be close to those submitted by President Reagan, since Democrats seem reluctant to ask for major increases in the science budget and Republicans are reluctant to demand major decreases.

According to a preliminary analysis of the Administration's budget request released last week by the American Association for the Advancement of Science (AAAS), the net result of the request would be to increase research and development spending by 2.2 per cent in real terms over 1981 to \$44,400 million.

Within this total, basic research will increase from \$5,100 million in 1981 to \$5,900 million in 1983, an increase of 14.9 per cent in current dollars, but a decrease of 1.3 per cent in constant dollars. The AAAS analysis demonstrates clearly that the expansion of research budgets is almost entirely explained by the growth of military-supported research.

Much grouching has been heard on Capitol Hill over the past few weeks as different committees have held hearings on the Administration's proposed budget. Some of the loudest have come from members of the space science community where, although major projects such as the Galileo mission to Jupiter have retained support in an overall budget that increases by 28 per cent between 1981 and 1983, several research programmes will still have to be terminated.

Quoting examples such as the decision to close the lunar curatorial facility, the main centre for the analysis of Moon samples and cosmic dust collected by the Apollo missions, as well as to shut down the transmission and reception of data from Pioneers 10 and 11, launched in the mid-1970s and still sending data from deep space, the planetary sciences division of the American Astronomical Society has released a statement claiming that the Reagan budget would devastate planetary research at the National Aeronautics and Space Agency.

A strong critic of the Administration's proposals for energy research has been Dr John Deutch, who was director of the Department of Energy's Office of Energy Research for the first two years of the Carter Administration, and is now dean of science at the Massachusetts Institute of Technology. Dr Deutch told a Senate subcommittee last week that the proposed cuts in energy research and development supported by the federal government would have a damaging effect on the nation's future energy supply. The private sector could not be expected to undertake all necessary research and development in energy, said Dr Deutch: "a policy based on this assumption places our technical future in energy technology at risk".

Whereas the Administration is planning to cut support for coal research and

development from \$500 million in 1982 to \$84 million next year, Dr Deutch suggested it should be raised to \$625 million. In contrast, rather than asking for \$557 million for the support of breeder reactor systems research, he suggested cutting this back by \$125 million and abandoning the proposed liquid metal fast breeder reactor at Clinch River in Tennessee, which he described as "unneeded and technically obsolete".

The most important factor determining the final outcome of the science budget will be the broader political debate on how much of a federal deficit the nation's economy can tolerate. If Congress decides that the proposed deficit is too large, but is reluctant to cut back on demands for increased military spending (or to increase tax revenues), then it will be looking for more cuts in the civilian budget. Whether this will be necessary depends on deliberations within the budget committees of the Senate and the House of Representatives that will be taking place over the next two months.

David Dickson

Glenn speaks out

One of the strongest critics of the Administration's policies towards the support of research and development is Senator John Glenn of Ohio who, like Senator Harrison (Jack) Schmitt, the chairman of the Commerce Committee's science and technology subcommittee, was a member of NASA's early astronaut team.

Senator Glenn has been addressing scientific and engineering societies, in



Glenn — from spaceship to White House?

particular advocating strong unilateral measures to prevent the proliferation of nuclear weapons through controls on nuclear exports. He is also expected to be a candidate for the Democrat presidential nomination in 1984, offering an alternative vision of liberalism to those of Senator Edward Kennedy and former vice-president Walter Mondale.

CORRESPONDENCE

Professional pride

SIR — On page 542 of your issue of 18 February you inveigh against the "deceit" of groups of professional people who combine to campaign against nuclear weapons. On page 545 we read "... the Pentagon's Defense Science Board, chaired by Dr John Deutsch of the Massachusetts Institute of Technology, recommended a start on the production of binary weapons and that the Department of Defense should prepare for a major increase in its chemical warfare programmes". I do not recall having seen in your editorial columns any denunciation of *this* kind of professional activity. Apparently it is all right for groups of bought scientists, operating mainly in secret, to promote the development of the instruments of mass murder, but somehow deceitful for groups of free scientists, meeting in public, to take the contrary stance.

J.R.S. FINCHAM

Department of Genetics,
University of Edinburgh,
Edinburgh, UK

Protein evolution

SIR — Hoyle and Wickramasinghe believe, probably correctly, that the age of the Earth is not great enough to have given time for chance mutations to produce enzymes specified at every point in a chain of, say, 200 amino acid units. Jukes (*Nature* 18 February 1982, p.548) replies that proteins evolve by elongation of small polypeptides. This reply is insufficient if the elongation process must still produce a unique arrangement of amino acid units of the chain length of the final enzyme.

The size of the initial polypeptide is irrelevant. The essential point is (see also *Nature* 3 December 1981, p.396) that enzymatic activity of a particular kind can arise from quite short lengths of polypeptide chain, for example six amino acid units. The rest of the chain, if present, may gain significance in the course of evolution, and must certainly have the property of not folding so as to enclose the active part of the chain. The presence of life on this planet does not, however, depend, as Hoyle supposes, on the emergence of the complete chain, specified point by point. The minimal length required for a particular enzymatic activity would not require Hoyle's endless aeons to appear by chance; it would be already present in the primitive ocean.

A.E. ROUT

British National Oil Corporation
(Trading) Ltd,
London W1, UK

A wing and a prayer

SIR — Is Jukes a secret believer, or merely an unwary fox in among the fundamentalist chickens? The ancestral relationship he implies between the biplane and the Boeing 747 (*Nature* 18 February, p.548) is one of design, not of descent. This is a striking counterpart of the creation process. Tragically, God's Concorde has fallen a long way since his maiden bite at the evolutionary tree, and even American jumbos finally land up in (we hope) US junkyards. Pigs might, but they have the good sense not to try.

C. DARNBROUGH

The University of Glasgow,
Glasgow, UK

Electricity costs

SIR — Professor Fremlin (*Nature* 1 October 1981, p.332) wonders why the retail price index is used to correct for inflation. The short answer is given by the Central Electricity Generating Board's (CEGB's) 1980-81 report, paragraph 168. "The cost per kWh of fuel ... rose by 18.6 per cent (between 1979-80 and 1980-81). This was higher than the general inflation rate (a 16.3 per cent increase in the RPI for a number of reasons". RPI is used almost universally as a measure of the general inflation rate. It also favours nuclear costs.

I have previously dealt¹ fully with the question of interest rates for comparison purposes, and if Professor Fremlin wants to discover an investment giving a steady 5 per cent above inflation rate I suggest he keeps an eye on his electricity bills, because they will have to cover a real 5 per cent interest on the funds being built up for decommissioning etc., by CEGB.

As far as accuracy is concerned, CEGB cost figures are given to 1 in 10,000, so my 1 per cent accuracy in the tables of my letter (see *Nature* 27 August 1981, p.791) hardly requires a probable error. Where "estimated minimum (or probable) corrections" are given, the "error" arises from the refusal to release available accurate information and its magnitude is known only to CEGB.

On future action, until Professor Fremlin is prepared to take real account of the work of Gerald Leach and his colleagues², I cannot take seriously such *ex cathedra* statements as his last sentence, "As of now it (nuclear power) is the only means capable of the scale of output needed".

I accept that cost ratios in 25 years' time are far more uncertain than the differences between my figures and those of CEGB. (Although it seems that there is considerably more uncertainty on nuclear than on coal costs.) The importance of the calculations of present costs is that CEGB claimed to be able to produce cheaper electricity by building new nuclear power stations immediately (Heysham II and the Sizewell PWR) even though they were not needed to satisfy demand. It has now been demonstrated¹ that this claim was entirely dependent on the assumption of a massive (36 per cent) increase in the real cost of coal to CEGB over the six years of the construction period (1980-86). This assumption was very effectively camouflaged in the board's documents and this is not surprising since it was made at almost exactly the same time as the "understanding" was entered into with the National Coal Board which practically guarantees that the real cost of coal to CEGB will not rise at all between 1980 and 1985.

In that case building Heysham II will certainly result in dearer electricity than if it had not been built, but instead the large overcapacity of coal-fired stations, which are at present being prematurely retired, had continued in operation.

Building unnecessary expensive new power stations may not bankrupt CEGB, but it will not help industry get cheaper electricity.

J.W. JEFFERY

Birkbeck College,
University of London, UK

1. Jeffery, J.W. *Energy Policy* 10, No. 2 (in the press).
2. Leach, G. et al. *IIED Sci. Rev.* (1979).

PCBs in rice oil

SIR — In March 1979, patients suffering from chloracne began appearing on the west coast of Taiwan. By October the disease had reached epidemic proportions, affecting over 1,100 people in the T'aichung and Changhua prefectures.

The exposure factor common to these patients was the consumption of rice oil produced by the same rice oil processing company situated in Changhua prefecture. Investigation of the manufacturing process identified a leak in a heat exchanger, which resulted in oil contamination with Kanechlor 400, a 48 per cent chlorinated biphenyl. In addition to symptoms associated with chloracne, clinical signs included respiratory symptoms, decrease in conduction velocity in peripheral sensory nerves, diminished growth in young children, and four reported cases of pigmented skin in newborns, all of whom died within a year of birth.

The same type of mass rice oil contamination was reported in western Japan in 1968¹, affecting more than one thousand victims. These were cases of direct consumption of high doses of polychlorinated biphenyls. The use of these substances was banned in Taiwan in 1972, but that did not prevent this recurrence, which is evidence that Kanechlor 400 is still being used in the heat exchange process during the production of oil for human consumption. This recurrence should serve as a warning to those Asiatic rice-producing countries which still use these oil production systems.

WILLIAM Y.B. CHANG

University of Michigan,
Ann Arbor, Michigan, USA

1. World Health Organization *Envir. Hlth Criteria* 2 (1976).

Holmes' lives on

SIR — What's this nonsense about *Holmes' Principles of Physical Geology* being "long out of print"? ("What Makes a Good Textbook"; *Nature* 11 February, p.459.) This work has never been out of print. A comprehensively revised third edition was prepared by Sir Arthur's widow Doris (herself a fine geologist in her own right) and published in a revised format in 1978. The book has, as you say, been a great success both critically and commercially. It continues to be so after nearly 40 years of *uninterrupted* availability.

DOMINIC RECALDIN

Thomas Nelson & Sons Ltd,
Walton-on-Thames, Surrey, UK

An end to creation?

SIR — Why do you continue to publish letters on creationism? I am reasonably confident that you would not publish arguments on whether the Earth is flat or whether epilepsy is caused by demoniacal possession.

Of course, scientists are entitled to their private superstitions; but they have no place in a supposedly serious journal. Pray let us see those long overdue words: "This correspondence is now closed."

M. HAMMERTON

Department of Psychology,
University of Newcastle upon Tyne, UK

Science on the Space Shuttle

W.M. Neupert, P.M. Banks, G.E. Brueckner, E.G. Chipman, J. Cowles, J.A.M. McDonnell, R. Novick, S. Ollendorf, S.D. Shawhan, J.J. Triolo & J.L. Weinberg

THE third flight of the Space Shuttle, chiefly intended to test the orbiting portion of the Shuttle in extreme thermal conditions in space, nevertheless carries a scientific payload. The investigations will demonstrate the Shuttle's capability for research in space plasma physics, solar physics, astronomy, life sciences and technology and will also determine the effects that the presence of the Orbiter may have on its immediate environment. The information to be gathered is thus the foundation for planning of future investigations with the Space Shuttle.

The scientific payload is designated OSS-1 because the programme was originally managed by the Office of Space Science (now the Office of Space Science and Applications) at NASA headquarters. Responsibility for development of the payload was assigned to the Goddard Space Flight Center. The manager for the scientific payload is Kenneth Kissin. The OSS-1 instruments will study the Orbiter's plasma environment and the propagation of an electron beam in space, record the Sun's UV and X-ray fluxes, observe the zodiacal light and Milky Way, record interplanetary dust particle impacts, operate a sophisticated heat pipe system and also grow plants in the Orbiter's cabin. Demonstrating the Orbiter's research capability demands accurate control of the Orbiter's orientation during solar observations and use of a Remote Manipulator System to carry a package of sensors that will map charged particle and

magnetic and electric field distributions around the vehicle.

The same instruments will make sensitive measurements of how the Orbiter alters its environments by the emission of particles or electromagnetic fields. Although the spacecraft has been designed to minimize such effects some, such as propellant plumes from thruster engine firings or electromagnetic radiation from the electrical circuits, are unavoidable.

The nine instrument packages on the third Shuttle flight should yield important data for space plasma physicists, astronomers, life scientists and engineers and also pave the way for future scientific payloads.

All but one of the OSS-1 instruments are mounted on an engineering model of the Spacelab pallet manufactured by the European Space Agency (Fig. 1). The instruments, pallet and the various subsystems for command, data handling, power and thermal control weigh 3,132 kg. The Plant Lignification Experiment is located in mid-deck lockers of the Orbiter cabin and two OSS-1 tape recorders and two control panels are located in the aft flight deck on the upper level of the cabin. Most internal experiment operations will be carried out via commands from the investigators located at the Payload Operations Control Center at Johnson Space Center in Houston, Texas.

On the third Shuttle flight, scheduled for late March, the Orbiter will be placed in a circular orbit of 241 km with an inclination of 38°. During the seven-day flight, the Orbiter will be held in several attitudes as part of the mission's thermal test objectives. The 28-hour long orientation of the Orbiter bay towards the Sun is suitable for solar observations. The planned 80-hour long period when the Orbiter will be maintained with its nose towards the Sun is a prime observing interval for the plasma physics investigations.

The nine experiments will now be described separately.

Electromagnetic environment

During the next decade, active electron and ion beam experiments in space will become an important tool for probing the environment above the Earth's atmosphere. This region contains plasma and magnetic fields of both solar and terrestrial origin. Hitherto, actively controlled charged particle experiments carried by rockets and unmanned satellites have been used to probe the structure of the Earth's magnetic field and to create the

physical conditions that led to the formation of the aurorae, seen at high latitudes¹.

The Vehicle Charging and Potential Experiment is the first opportunity to make vehicle charging and electron beam studies from a manned spacecraft. In the first instance, measurements will be made of the effect of the natural ionospheric environment on the gross electrical characteristics of the Orbiter using two pairs of charge and current probes, situated on opposite corners of the pallet, which simulate both the electrically insulating and conducting portions of the Orbiter's surface, and with a Langmuir Probe-Spherical Retarding Potential Analyser located on the sill of the pallet.

Active electron emissions will be used to change the natural electrical balance to

determine how a charged Orbiter will affect electron beam and direct plasma measurements. The electron source (the Fast Pulse Electron Generator) emits a 100-mA beam of nearly monoenergetic 1 keV electrons. Pulses as short as 600 ns or as long as 109 s can be generated (Fig. 2a).

The Fast Pulse Electron Generator will also be used to investigate the effectiveness

Table 1 OSS-1 Plasma Diagnostics Package instrumentation and measurement ranges

● Low energy proton and electron differential energy analyser Nonthermal electron and ion energy spectra and pitch angle distributions for particle energies between 2 and 50 keV ● A.c. magnetic wave search coil sensor Magnetic fields with a frequency range of 10–30 kHz ● Total energetic electron fluxmeter Electron fluxes between 10^9 and 10^{14} electrons $\text{cm}^{-2} \text{ s}^{-1}$ ● A.c. electric and electrostatic wave analysers Electric fields with a frequency range of 10 Hz–1 GHz S-band field strength meter ● D.c. electrostatic double probe with spherical sensors Electric fields in one axis from 2 mV m^{-1} to 2 V m^{-1} ● D.c. triaxial fluxgate magnetometer Magnetic fields from 12 mG to 1.5 G ● Langmuir probe Thermal electron densities between 10^4 and 10^7 cm^{-3} Density irregularities with 10 m–10 km scale size ● Retarding potential analyser/Differential velocity probe Ion number density from 10^2 to 10^7 cm^{-3} Energy distribution function below 16 eV Directed ion velocities up to 15 km s^{-1} ● Ion mass spectrometer Mass ranges of 1–64 AMU Ion densities from 20 to $2 \times 10^7 \text{ ions cm}^{-3}$ ● Pressure gauge Ambient pressure from 10^{-3} to 10^{-7} torr

Werner Neupert is OSS-1 Mission Scientist and Head of the Solar Plasmas Branch, Laboratory for Astronomy and Solar Physics, at NASA Goddard Space Flight Center, Greenbelt, Maryland; Peter Banks, Principal Investigator of the Vehicle Charging and Potential Experiment, is Adjunct Professor Physics at Utah State University, Logan, Utah, and Professor of Electrical Engineering at Stanford University, Stanford, California; Guenter Brueckner, Principal Investigator of the Solar Ultraviolet Irradiance Monitor, is Head of the Solar Physics Branch at the Naval Research Laboratory, Washington DC; Eric Chipman is Program Scientist for the OSS-1 mission at NASA Headquarters, Washington DC; Joe Cowles, Principal Investigator of the Plant Lignification Experiment is Professor of Biology, University of Houston, Houston, Texas; J.A.M. McDonnell is Principal Investigator of the Microfoil Abrasion Experiment and Reader in Space Sciences at the University of Kent, Canterbury, UK; Robert Novick, Principal Investigator of the Solar Flare X-ray Polarimeter, is Professor of Physics, Columbia University, New York; Stanford Ollendorf, Principal Investigator of the Thermal Canister Experiment is Head, Spacecraft Component Development and Analysis Section, Goddard Space Flight Center; Stanley Shawhan, Principal Investigator of the Plasma Diagnostics Package, is Professor of Physics, University of Iowa, Iowa City, Iowa; Jack Triolo, Principal Investigator of the Contamination Monitor Package is in the Spacecraft Component Development and Analysis Section, Goddard Space Flight Center, Greenbelt, Maryland; and J.L. Weinberg, Principal Investigator of the Shuttle Spacelab Induced Atmospheres Experiment, is Research Scientist and Director of the Space Astronomy Laboratory at the University of Florida, Gainesville, Florida.

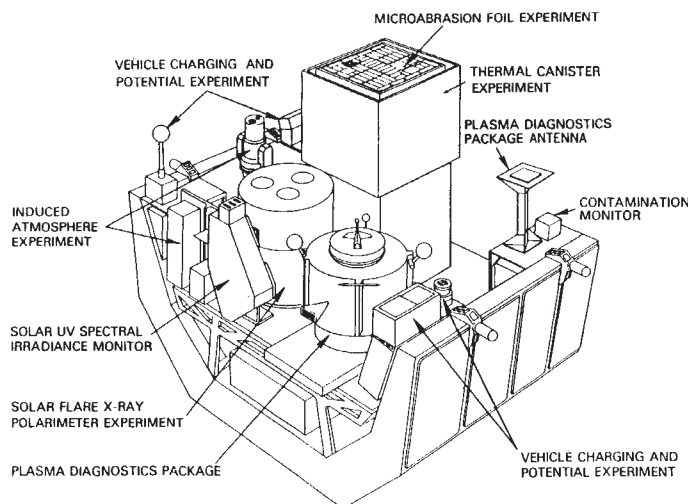


Fig. 1 OSS-1 instrumentation mounted on an engineering model pallet supplied by the European Space Agency. An additional OSS-1 experiment is carried in the mid-deck section of the Orbiter cabin.

of generating electromagnetic radiation by pulsing the electron beam at frequencies as high as 1 MHz. Several ground-based observatories will attempt to receive this radiation. Of particular interest is the extent to which VLF radiation can be generated using special electron beam modulation sequences. Detection will be attempted both at ground observatories and with VLF receivers aboard the recently launched Dynamics Explorer satellites.

Evidence will also be sought for a laboratory phenomenon termed beam-plasma discharge, in which an electron beam energizes a plasma column which then emits intense optical radiation and radio waves (Fig. 2a). It is thought that the electron beam initially ionizes residual gas, producing a columnar plasma. As the plasma density builds up, the beam creates electrostatic plasma waves which impart additional energy to the plasma, leading to further ionization and excitation and causing the column to emit optical and radio waves. In Earth-orbit, the different conditions may mean that beam plasma discharge does not occur — the gas density is less, the beam is not terminated by the metallic top of a chamber and the Orbiter moves the beam across magnetic field lines so that plasma may not build up. The Vehicle Charging and Potential Experiment will be used with the Plasma Diagnostics Package to tell whether beam plasma discharge is initiated by the beam from the orbiting electron generator (Fig. 2b).

Ionosphere studies

The Earth's ionosphere can be studied by introducing perturbations such as chemical tracers, radio waves and particle beams and by creating plasma wakes around solid bodies. The objectives of the Plasma Diagnostics Package, an assembly of electromagnetic and particle sensors, are to assess the electromagnetic and plasma environment of the Orbiter, to study the interaction of the vehicle with the surrounding plasma, to test the capabilities

of the Remote Manipulator System and to carry out active beam-plasma experiments in conjunction with the Fast Pulse Electron Generator. Measurements will be made of electric and magnetic fields, plasma waves, energetic ions and electrons and plasma parameters — density, composition, temperature and directed velocity (see Table 1).

The Plasma Diagnostics Package will be operated both on the OSS-1 pallet and while deployed by the Remote Manipulator System. As the Plasma Diagnostics Package is moved in and around the Orbiter bay, measurements will be made of the ambient pressure and of the spectrum of electromagnetic interference generated by the Orbiter's electrical subsystems. The pressure profiles in time and in distance from the Orbiter are relevant for the design of instruments sensitive to gaseous contamination and those requiring low operating pressures. The sensitivity of wave receivers and of topside ionospheric sounders to be flown on future Spacelabs will be determined by the levels now measured, while measurements of electric fields and particles by the Plasma Diagnostics Package will provide an independent assessment of the charge condition of the Orbiter.

The Orbiter will move in the ionosphere, at supersonic velocity (Mach number 6 relative to the characteristic plasma sound speed or the ion acoustic sound speed²) and so will create a plasma wake that may be identified by plasma depletion, energization of particles and the creation of Alfvén waves behind the Orbiter. Such processes are thought to be important consequences of the motion of other bodies through plasmas — for example, Alfvén waves behind the jovian moon Io may accelerate the particles which cause decametric radio noise bursts³. Remote Manipulator System trajectories have been designed to move the Plasma Diagnostics Package through the wake boundary, thus providing direct observations of its physical properties. When the package

flies again on the Spacelab-2 mission as a subsatellite, the wake will be examined out to 20 km behind the Orbiter⁴.

The combination of the Fast Pulse Electron Generator and the Plasma Diagnostics Package provides an opportunity to study the interactions of a beam of accelerated electrons with the ambient space environment (see Fig. 2b). Radio waves over a wide frequency range may be stimulated by both pulsed and continuous operation of the electron beam and measured with the Plasma Diagnostics Package. The two instruments will also operate together to investigate beam-plasma interactions such as the beam plasma discharge. Similar experiments will be conducted with more intense beams on Spacelab-1 and on later missions⁵.

Solar ultraviolet radiation

Solar UV radiation in the spectral range 120–300 nm has an important role in the energy balance and photochemistry of the Earth's upper atmosphere. Molecular oxygen absorbs radiation at wavelengths below 242 nm, resulting in its dissociation into atomic oxygen; ozone is dissociated into molecular and atomic oxygen by UV radiation below 310 nm. These reactions

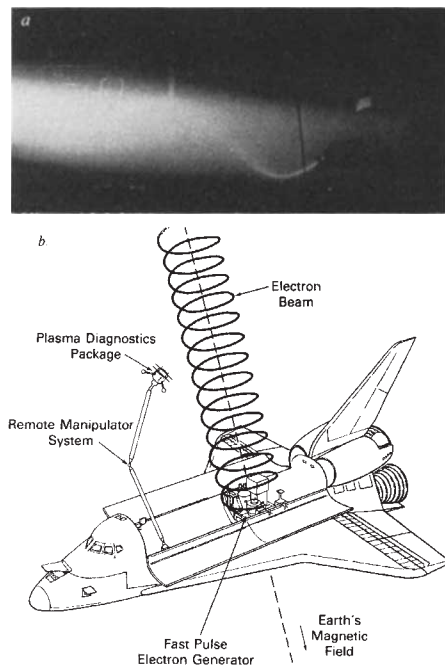


Fig. 2a, Photograph of light emission resulting from the interaction of electrons generated by Fast Pulse Electron Generator with residual atmosphere in a space simulation chamber at the Johnson Space Center. The column of luminosity enclosing the helical path of the primary electron beam is due to the occurrence of a beam plasma discharge for the conditions pertaining to this particular operation of the electron generator. **b**, Scheme for the joint Fast Pulse Electron Generator and Plasma Diagnostics Package operations. As the electron beam is emitted along some angle to the Earth's magnetic field, the Remote Manipulator System sweeps the Plasma Diagnostics Package back and forth across the beam region to make measurements of plasma fields and waves and of the energy distributions of electrons and ions.

take place at altitudes from 30 to 160 km in the Earth's atmosphere. Ozone is formed from molecular and atomic oxygen by a catalytic reaction which involves trace species such as NO, NO_x or Cl. Quantitative description of these reactions requires precise knowledge of the absolute amount of solar UV radiation as a function of wavelength. Although many attempts have been made to measure this quantity, large experimental discrepancies still exist⁶ (Fig. 3). The objective of this first flight of the Solar Ultraviolet Spectral Irradiance Monitor experiment is to establish a more accurate base of solar UV irradiance measurements with an absolute error of 10 per cent or less over the wavelength region 120–400 nm. During the STS-3 flight, it will accumulate approximately 20 hours of solar measurements, compared with 5 minutes during a typical sounding rocket flight.

Calibration is important. The instrument carries two independent spectrometers and an in-flight calibration light source which allows tracking of any sensitivity change due to vibration at lift-off or contamination during flight. Its seven detectors will allow cross-checks of possible detector changes. It can be operated in broadband (5 nm) or narrow band (0.15 nm) modes over the wavelength region 120–400 nm. During 17 orbits amounting to approximately 28 hours, the bay of the Orbiter will be pointed to the Sun by the crew, using Sun sensors mounted on the Solar Ultraviolet Irradiance Monitor and measurements of solar intensities will be interleaved with periodic in-flight calibrations⁷.

The measurement of the solar UV irradiance on OSS-1 is only the first step of a programme to measure the variability of the solar UV radiation over an 11-year solar cycle. This variability has been estimated to be less than 20 per cent in the 170–210 nm region and less than 2 per cent in the 210–300 nm region⁸, and Fig. 3 shows that a decisive improvement of accuracy is needed if the Sun's UV variability is to be determined.

Solar flares

Although it is known that there are high-energy electrons in the Sun's atmosphere during solar flares, fundamental questions remain about the nature and location of the

mechanisms by which the electrons are energized and their energy dissipated. It is, however, believed that a flux of magnetic energy from the solar interior provides the energy for these events, and that interactions of the accelerated particles with the Sun's atmosphere dissipate this energy, leading to observable flare phenomena including X-ray emission. Similar bursts of energy observed by the Einstein Observatory from other stellar objects have established that flares are relatively common. Hard X rays emitted during flares carry unique information about the motion of the electrons producing the radiation⁹. A definitive observation of the state of polarization of the radiation could provide important data to test theoretical models of the flare phenomenon.

The Solar Flare X-ray Polarimeter on OSS-1 aims to observe flare X rays emitted between 5 and 30 keV and to measure their polarization as a function of time and photon energy. The instrument uses blocks of metallic lithium surrounded by xenon-filled proportional counters as detectors. If polarized, the incident X rays will be scattered preferentially by the lithium into directions normal to the plane of the electric vector of the incoming radiation. To avoid instrumental effects that have plagued previous measurements, the instrument uses three independent sets of scattering blocks and detectors, with each unit rotated by 120° with respect to the other two about a line passing through the Sun. A minimum of two units is necessary to determine the magnitude and orientation of polarization; the use of a third provides redundancy and increased effective area¹⁰.

The instrument will be aimed at the Sun by orienting the entire bay of the Orbiter and it is planned to maintain the Orbiter in this orientation for approximately 28 hours. Solar flares occur only sporadically on the Sun, so that the observation of flare emission is not assured. However, the instrument has sufficient sensitivity that even a small event can yield a usable signal. Such events can be expected to occur about once a day on average.

Zodiacal light

The zodiacal light arises from sunlight scattered or absorbed by interplanetary dust particles, the characteristics of which have been partly defined by rockets and unmanned Earth-orbiting satellites¹¹. Observations of colour, polarization and angular dependence are needed to determine dust particle size and composition¹². Unfortunately, accurate data are scarce. Thus, polarization results at different wavelengths are often combined. There are some observations of brightness, but few of polarization and colour in regions off the ecliptic, closer than 30° to the Sun and near the anti-solar point — the regions which contain the most information on the dust. The Shuttle Spacelab Induced Atmosphere Experiment

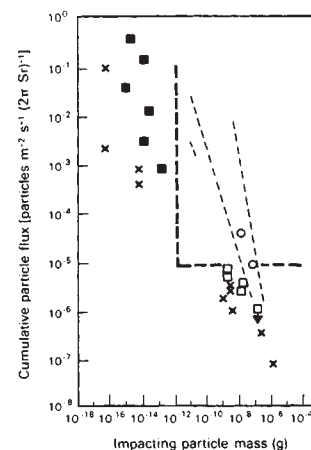


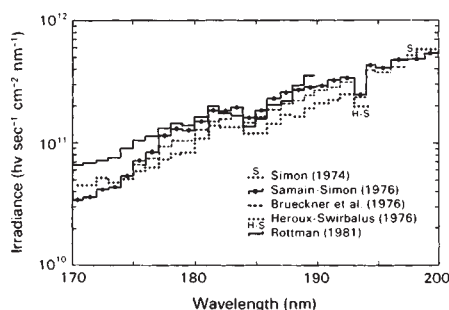
Fig. 4 Measurements of near-Earth microparticle fluxes by various techniques: ■, impact plasma sensors; ×, capacitor discharge sensors; ○, recovered surface examination; □, perforation of pressurized cylinders; dashed lines, models derived from microphone sensors. The area to the upper right of the heavy dashed line is the sampling regime of the Microabrasion Foil Experiment on OSS-1. The cumulative particle flux is the flux of particles above a specific mass. Adapted from ref. 14.

will provide observations to determine more precisely the position of maximum concentration of interplanetary dust and will search for evidence of different particle characteristics for dust near the Sun.

Although astronomical observations from space avoid the effects of the Earth's atmosphere, the observing platform releases particulates which may form a local contaminant cloud that will also scatter sunlight into the detector. The presence of such an induced atmosphere can be established by comparing observations made when the spacecraft is in sunlight and in the Earth's shadow. The photopolarimeter system on OSS-1 will measure the light scattered from any local cloud of particulates and will record the brightness, polarization, colour and angular dependence of the diffuse astronomical background (zodiacal light and background starlight) at visual and near IR wavelengths.

The instrumentation to be used is the spare unit of the Skylab S073 photopolarimeter and bore-sighted 16 mm camera. The system, which includes an optical train, optical filters for 10 wavelength bands between 400 and 820 nm, a polarization analyser, a brightness calibration source and a photoelectric sensor, can carry out observing sequences according to pre-programmed routines. Although the instrument was operated on Skylab, difficulties with the airlock precluded observations closer than 80° to the Sun. A new gimbal mount has been designed for this mission, allowing the instruments to be scanned in a vertical plane running fore and aft along the Orbiter axis. The instrument will view the entire sky between 20° and 120° from the Sun. Limited observations will also be possible at larger angles.

Fig. 3 Recent measurements of solar UV irradiance in the spectral region between 170 and 200 nm (after ref. 6).



Interplanetary dust

Direct sampling of interplanetary particulate matter has been carried out from balloons, U-2 airplanes, rockets, Earth-orbiting spacecraft¹³. Remote measurements have been made with sensors in Earth orbit and interplanetary space¹⁴. Even impact craters in lunar samples have provided valuable information. A wide range of particle masses, down to 10^{-15} g have been observed, with the cumulative particle flux increasing strongly with decreasing mass (Fig. 4). While comets are thought to be the most likely source of interplanetary dust, direct measurement of particle composition could discriminate between cometary and asteroidal sources for the dust.

The Microabrasion Foil Experiment aims to measure the high velocity microparticle flux in near-Earth orbit, for particle masses greater than 10^{-12} g, to investigate the density distributions of the impacting particles and to study their chemical properties by analysis of residues remaining in the impact craters. The sensor, attached to the thermal blanket on top of the thermal canister experiment, consists of approximately 1 m^2 of $5 \mu\text{m}$ thick aluminium foil bonded to a gold-coated brass support mesh bonded to a Kapton sheet to form a double layer detector.

Differing types of impacts may be recorded. A particle of mass $< 10^{-12}$ g will form a hypervelocity impact crater on the foil surface, without penetrating the foil. More massive particles will penetrate the foil forming a characteristic 'penetration profile', dependent on the particle's incident velocity. Particles penetrating the foil may survive intact to produce a single impact crater on the rear Kapton sheet, or the particle may split into fragments and produce a corresponding number of impact craters on the sheet. Low density 'fluffy' particles readily fragment

whereas high density iron or stony micrometeorites survive almost intact.

Unlike the other experiments, the Microabrasion Foil Experiment is passive. Post-flight measurements of the craters with scanning electron microscopy and energy dispersive X-ray microprobe analysis of residues will provide information on elemental composition, density and shape, and thus on the origin of these particles.

Plant lignification

Although few plants have been grown in space, Russian experiments have demonstrated that near-zero gravity disorients root and shoot growth, enhances plant sensitivity to substrate moisture conditions and generally results in a high mortality rate. However, little is known about the physiological changes that occur. Understanding of gravity's effects on plant growth and metabolism will provide an insight into plant physiology and aid development of an effective biological life support system.

After cellulose, lignin is the most abundant carbon compound in plants and provides both their strength and form. As gravity is believed to be a primary controlling stimulus for lignification the Plant Lignification Experiment will evaluate how near zero gravity affects the quantity and rate of lignin formation in different plant species during early stages of development.

Of the major groups of higher plants the gymnosperms will be represented by slash pine, the monocotyledonous angiosperm by oat, and the dicotyledonous angiosperm by the mung bean. All are compact species that may be grown in the limited space and relatively low light levels provided by the compact flight Plant Growth Unit. Oat and mung bean seeds and young pine seedlings will be planted in the growth chambers (Fig

5) before launch so that most seedling development will take place in a weightless environment. Electronics for controlling and monitoring temperature and light cycles are incorporated into the Plant Growth Unit. At the latest convenient time (about 7 hours before launch) the unit will be carried on board and installed in a mid-deck locker of the Orbiter cabin, where it will remain throughout the flight.

On landing, the Plant Growth Unit will immediately be removed from the Shuttle, the plants photographed, and the gaseous atmosphere of the plant chambers analysed. The seedlings will then be removed and analysed for lignin content.

Control experiments, with the plants growing in a 1-g environment, will be conducted after the flight using the flight hardware and flight environmental data. Lignin data from the flight and control plants will be compared for patterns of lignin deposition to assess whether lignin is reduced in plants grown in zero gravity.

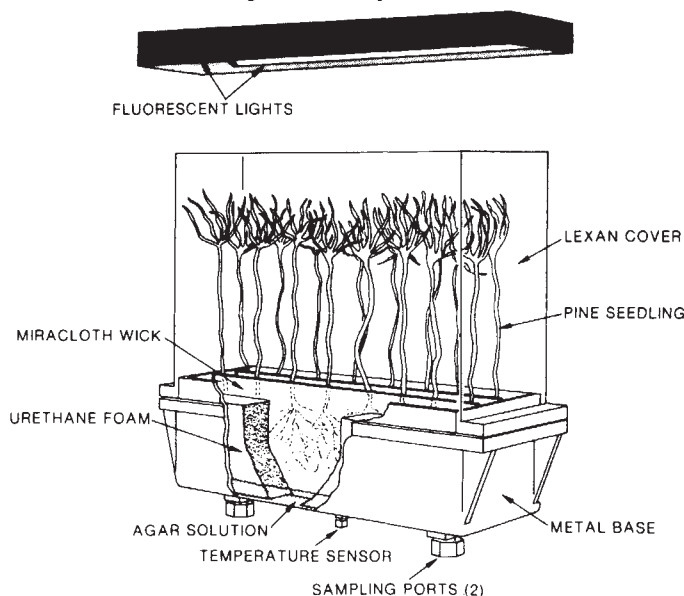
Because the reduced flight time of the previous Orbiter flight prevented completion of the Heflex (plant) Bioengineering Test that it carried, this experiment has been added to the current flight and is also carried¹⁶ in the Orbiter cabin. The test is being conducted specifically to determine the optimal soil moisture conditions for germination and growth of the sunflower *Helianthus annuus* in near zero gravity. It is a prelude to a plant growth experiment to be conducted on the first Spacelab mission to be flown on the Shuttle in 1983.

Thermal canister

The long-term use of Space Shuttle means that many scientific and technical investigations can be performed in the Orbiter bay. However, the extreme thermal environmental conditions ranging from equivalent sink temperatures of $+100^\circ\text{C}$ in full Sun, to -100°C , in shadow may cause problems. In the past such conditions were accommodated using coatings, insulation and heaters. With the Shuttle, an instrument designed for one set of conditions may have to survive in an entirely different environment if flown again with different orbit attitudes. If a thermal enclosure were provided which decoupled instruments from the wide extremes in external temperature whilst maintaining them in a benign environment, simpler thermal designs for instruments, with limited maintenance between flights, might be realized.

To this end the Thermal Canister Experiment aims to determine whether a device using controllable heat pipes could maintain simulated instruments at several selectable temperature levels in zero gravity, and under widely varying internal and external thermal loads. It is hoped to demonstrate $\pm 3^\circ\text{C}$ temperature stability at various control points in the canister while dissipating up to 400 W in cold Orbiter attitudes (bay away from the Sun)

Fig. 5 View of a growth chamber, one of six used in the Plant Growth Unit of the Plant Lignification Experiment.



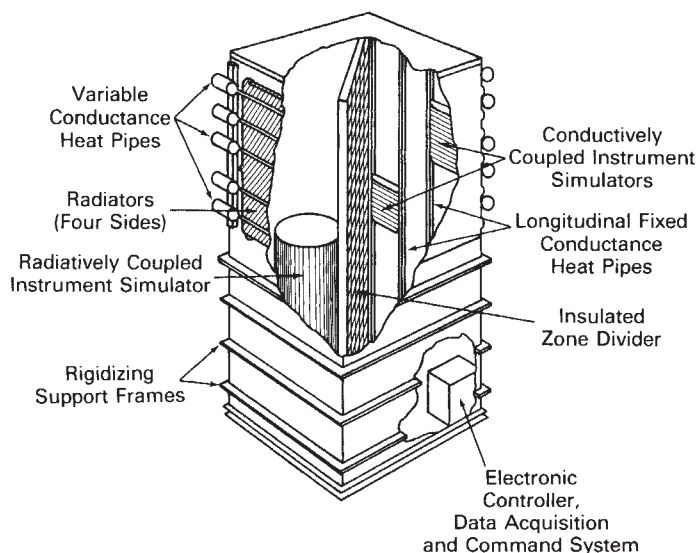


Fig. 6 Cutaway view of the Thermal Canister Experiment showing controllable heat pipes. In flight conditions, the lower half of the unit is covered with multi-layer insulation to provide a measure of isolation from temperature extremes in the Orbiter bay.

and 100 W hot (bay towards the Sun) conditions.

The Thermal Canister Experiment (Fig. 6) consists of a rectangular enclosure 3m high $\times$ 1m long $\times$ 1m wide with its aluminium sides equalized in temperature by a system of longitudinal fixed conductance heat pipes. These heat pipes collect the thermal energy dissipated internally by electrical heaters simulating instruments in operation and the energy absorbed from direct and reflected sunlight. This heat is then conducted to variable conductance heat pipes mounted to external radiators at the upper end of the canister and radiated to space. The heat pipes are long narrow closed chambers with internal capillary wicking which provides pumping action. The wick is saturated with a volatile liquid (ammonia) in equilibrium with its vapour. Heat transport is established by applying heat at one end (the evaporator) and providing cooling at the other end (the condenser) with the heat being transferred as latent heat of vaporization. The flow path is completed by capillary forces in the wick.

The variable conductance heat pipes are more complex than the fixed conductance type in that they contain a non-condensable gas (nitrogen) stored in a reservoir at the condenser end of each pipe. As the temperature of the evaporator end of the pipe falls a heating element raises the temperature of the reservoir, causing the gas to expand into the condenser, blocking the condenser region and effectively stopping heat pipe action. The length of condenser rendered inactive depends on the temperature level along the pipe. Conversely, with increasing evaporator-end temperature, the gas will recede into the reservoir making more active area of the radiators available for heat rejection to space. The signal for activating the reservoir heaters is supplied through a feedback loop consisting of a temperature control sensor and either a hardwire proportional controller or a computer-

driven controller. The sensors are attached to the canister side walls or on simulated instruments located in two different zones separated by an insulating barrier. The simulators are either radiatively or conductively coupled to the canister walls.

During the mission, it is planned to: (1) operate the canister over set points (5–25°C) located on the walls and on the simulators themselves; (2) change the internal dissipation (in the simulators); and (3) demonstrate control in maintaining the two zones at differing temperatures. The system can be operated by a proportional controller maintaining a specific temperature at one sensor, or by a microprocessor that uses all available data to maintain the overall temperature of the canister at some level, in balance with the environment, irrespective of the preselected set point.

Contamination monitoring

Payloads operating in the bay of the Orbiter will be exposed to a variable gaseous environment. In addition to outgassing of the vehicle and the payloads themselves, the operation of altitude control systems, the venting of relief valves and the dumping of water for thermal control of the vehicle all represent molecular sources that may affect sensitive instrumentation, particularly equipment at cryogenic temperatures. Measurements have been planned for the series of four orbital test flights using the Induced Environment Contamination Monitor provided by the Marshall Space Flight Center and located, on this flight, behind the OSS-1 pallet in the Orbiter bay. Particulate measurements will be provided by the Shuttle/Spacelab Induced Atmosphere Experiment. A molecular contamination monitor on the OSS-1 pallet provides information on molecular species around the OSS-1 instruments, supplementing measurements from the Induced Environment Contamination Monitor.

The Contamination Monitor Package, sponsored by the United States Air Force, contains four temperature-controlled Quartz Crystal Microbalances which measure the accreted mass of molecular fluxes. These microbalances are identical to those contained in the Induced Environment Contamination Monitor but their operation can be monitored and the data analysed during the flight. Their temperatures may be reset by command after initial results have been analysed to optimize the operation of the sensors.

The instrument will monitor the mass accretion of condensable volatile materials during ascent, on-orbit operations and descent. After the mission, the data on mass build-up will be correlated with payload activities, Orbiter operational events, performance of other OSS-1 instruments and Induced Environment Contamination Monitor results. The activation energy of the major species of the accreted materials can also be estimated.

While minimal impact is expected for most instruments that may fly on Shuttle, long-term exposures of particularly sensitive optical components or cryogenic surfaces to the Orbiter environment may require special precautions.

Future prospects

The OSS-1 instruments will point the way to future developments in the use of the Shuttle for space science investigations. Two of the OSS-1 instruments will fly again on the Spacelab-2 mission: the Plasma Diagnostics Package will be released from the Remote Manipulator System to become a subsatellite and make measurements of the spacecraft wake at distances up to 20 km behind the Orbiter, while the Solar Ultraviolet Irradiance Monitor will be mounted on a solar pointer to continue UV irradiance measurements through the solar cycle. The data from other experiments will be used to design large UV and IR telescopes and other instruments for future flights in the Orbiter. Ultimately, the Shuttle program will include discipline-dedicated flights and the development of space platforms serviced by the Shuttle.

1. Grandal, B. (ed.) *Artificial Particle Beams Utilized in Space Plasma Physics* (Plenum, New York, 1982).
2. Samir, U. & Stone, N.H. *Acta Astronaut.* **1**, 1901–1141 (1980).
3. Gurnett, D.A. & Goertz, C.K. *J. geophys. Res.* **86**, 717 (1981).
4. Shawhan, S.D., Burch, J.L. & Fredricks, R.W. *AIAA Paper 82-0085*, Florida (1982).
5. Raitt, W.J. *et al.* *AIAA Paper 82-0083*, Florida (1982).
6. Simon, P.C. *Solar Phys.* **74**, 273 (1981).
7. VanHoosier, M.E., Bartoe, J.-D.F., Brueckner, G.E., Prinz, D.K. & Cook, J.W. *Solar Phys.* **74**, 521 (1981).
8. Cook, J.W., Brueckner, G.E. & VanHoosier, M.E. *J. geophys. Res.* **85**, 2257 (1980).
9. Elwert, G. & Haug, E. *Solar Phys.* **20**, 413 (1971).
10. Lemen, J.R. *et al.* *Solar Phys.* (submitted).
11. Weinberg, J.L. in *Lecture Notes in Physics* Vol. 48. (eds Elsässer H. & Fechtig, H.) 3–18 (Springer, Berlin, 1976).
12. Giese, R.H., Hanner, M.S. & Leinert, C. *Planet Space Sci.* **21**, 2061 (1973).
13. Brownlee, D.E. in *Cosmic Dust* (ed. McDonnell, J.A.M.) 295–336 (Wiley, New York, 1978).
14. McDonnell, J.A.M. in *Cosmic Dust* (Wiley, New York, 1978).
15. Leinert, C. *Space Sci. Rev.* **18**, 281 (1975).
16. Taranick, J.V. & Settle, M. *Science* **214**, 619 (1981).

NEWS AND VIEWS

How life began

from Preston Cloud

THE prebiotic Earth, the beginnings of life and the evolution of biological complexity were the subjects of a small symposium* of 27 participants from Australia, Canada, the Peoples' Republic of China and the United States held recently at Jindabyne, New South Wales.

The now-large group of polyatomic organic molecules known from galactic space as a result of radioastronomy linked with spectroscopy extends the plausible choices available for the initial building blocks of life. Brown (Monash University) described contributions from (1) chemical evolution in the dark nebulae of the Galaxy; (2) chemical evolution in the pre-solar nebula followed by the arrival on Earth of its products in planetesimal debris; and (3) chemical evolution in Earth's primitive atmosphere and hydrosphere. The arrival in the early solar nebula of cold organic molecules from the interstellar stockpile appealed to Chang (NASA); had they been warmed on entry, ices would have melted, and a lot of aqueous chemical evolution might have occurred in restricted environments.

The heavily cratered surfaces of the solid planets and satellites were cited by Taylor (ANU) as evidence that planetesimal infall continued until about 3.8 Gyr (3.8×10^9 yr) ago, wiping out earlier records of earth history. Meanwhile Earth's core and mantle had formed, a conclusion drawn by Sun (CSIRO) from the constancy from 3.8 Gyr ago until the present of the ratio of P_2O_5 to TiO_2 in mantle-derived basalts. The motions of a fluid core can generate the magnetosphere that shields Earth's surface from the potentially lethal solar wind.

With a source of prebiotic organic molecules and such shielding, chemical evolution along the continuum from non-living to living could occur at sites shielded from lethal UV wavelengths. The central problems of biopoesis (N.W. Pirie's word for the origin of life) are how to achieve the necessary concentrations of organic macromolecules, how to catalyse the essential reactions and how to accomplish chirality.

The chemistry of the atmosphere under

which life arose is important. Earlier views invoking highly reducing conditions are giving way to geochemical evidence for a nearly neutral atmosphere. Fluid inclusions in 3.5 Gyr old barite show CO_2 , H_2O , H_2S and possible hydrocarbons (Groves, UWA). An atmosphere at this time of CO_2 , H_2O , N_2 and CO , with minor proportions of NH_3 , H , CH_4 and H_2S , seems likely. Experiments on chemical evolution show small but finite yields of interesting organic molecules under such an atmosphere in the absence of free O_2 . Carbonaceous chondrites might even more successfully concentrate amino acids and nucleotide bases in wet clay banks that could promote dehydration condensations. And the chemistry of Archaean sediments argues for very low and evanescent free O_2 . In fact, the real problem with oxygen is not to account for its absence in earlier times but for its presence later. Because the co-products of the processes that produce it are all thermodynamically unstable, free O_2 can accumulate only as the product of a kinetic lag.

Sedimentary rocks 3.8 Gyr old are evidence for a substantial atmosphere and hydrosphere in earliest Archaean time. Despite the faint early sun of solar physicists, liquid H_2O has been present at Earth's surface from then onwards. McCulloch (ANU) described these oldest known sediments as implying derivation from undifferentiated source rocks — hence from crust formed no more than 200–300 million years earlier, about 4–4.1 aeons ago.

Returning to the question of biopoesis, the origin of replication was the subject of spirited discussion. Reaney (La Trobe University) proposed that RNA came first, as implied by the appearance of ribose before deoxyribose in existing biosynthetic pathways. But Langridge (CSIRO) reasoned that DNA came first because its bonds resist hydrolysis more strongly than those of RNA. Or did something else start the process? Burgess (Ludwig Institute) proposed that sugars were the logical beginners because they form readily, have a limited rotational freedom, will form helices with catalytic properties, are stable and link with lipids to form micelles.

But, in prebiotic evolution, how are sequences replicated and copying errors limited? Replication would be favoured, Reaney supposed, by day-to-night temperature variations. Experimentally, strands of DNA separate on heating and recombine when cooled. Such a mechanism could be important on an early Earth with shorter days and hence more rapidly repeated cycles.

Chirality may not be as much of a problem as once supposed. Brown remarked that once as many as five amino acids of the same handedness are strung together, they will preferentially add other amino acids of similar handedness from racemic mixes. To get five amino acids or sugars of the same handedness lined up by chance is entirely probable. And given a metabolizing, self-replicating structure, mutation is the next essential property. Copying error prescribes mutation and mutation predisposes to evolution.

Anderson (CSIRO) then showed how colloidal systems, bilayer vesicles and, ultimately, cells could result from the propensity for self-organization of amphipathic lipids and proteins. The distinctive properties of membranes permit the energy transduction that results in living cells.

Much attention at the symposium was directed to the earliest known forms of life and to their probable biochemical evolution. Evolution beyond fermentation requires photosynthesis and phosphorylation to trap and maintain a free-energy flow. By what routes, Anderson (CSIRO) asked, could photosynthesis and respiration arise? Mauzerall (Rockefeller University) sees biosynthetic pathways as recapitulating evolution. Each porphyrin on the path to chlorophyll biosynthesis had a specific function in its own time. The first porphyrins of the haem/chlorophyll biosynthetic pathway, the uroporphyrins, were presumably the initial photo-oxidants that later became coupled to phosphorylation and photosynthesis. Bacteriorhodopsin, present in halophilic bacteria, is the most primitive surviving

Preston Cloud is at the Baas-Becking Geobiological Laboratory, Box 378, Canberra City, ACT, Australia.

*The symposium was held from 13 to 15 December 1981 under the joint sponsorship of the Australian Academy of Science and the National Academy of Sciences of the USA.

molecular device to promote ion flow across membranes.

Walter (Baas Becking Laboratory) reviewed the fossil evidence and Hayes the biogeochemical evidence for early evolution. They concluded that life was present by 3.5 and photosynthesis by 2.8 Gyr ago in the Warrawoona and Fortescue strata of Western Australia.

An extended discussion followed about the importance of sulphur in living systems, the geological record of sedimentary sulphate and S isotopes, the possible importance of methanogens and the question of when free O₂ from photosynthesis first became an important atmospheric gas. Groves noted that sulphates are uncommon in Archaean rocks whereas sulphides are abundant. Donnelly added that the sulphur isotope record of stratabound volcanogenic sulphide deposits calls for very low concentrations of oxidized sulphur in the Archaean hydrosphere. The record observed is consistent with the convergence of three important and perhaps related events ~ 2 Gyr ago: (1) the beginning of free O₂ as a permanent component of the atmosphere, (2) the end of methanogenesis as an important function of the biomass, and (3) the onset of dissimilatory sulphate reduction.

The emergence of the eukaryotic cell at some time after ~ 2 Gyr ago was the focus of the final session at Jindabyne. Walter reviewed the biogeological evidence for the rise of O₂ during Earth history, the probable time of emergence of the eukaryotes and alternative hypotheses for eukaryotic origins. Runnegar (New England University) observed that even anaerobic eukaryotes need O₂ for the production of collagen and components of the cell wall. Thus eukaryotes probably made their first appearance later than 2 Gyr ago. But how? Langridge observed that differences between organellar and nuclear DNA of eukaryotes could result either by derivation of the organellar DNA from the nucleus by a transfer analogous to that involved in the origin of some viruses, or by endosymbiosis.

Lastly the conference turned to the origin of membrane-bound organelles, considered distinctive of the eukaryotic cell. Langridge noted that cytochrome c amino acid sequences and 5S rRNA nucleotide sequences imply a date of ~ 1.8 Gyr ago for the divergence of the eukaryotic from the prokaryotic level of organization — a date consistent with biogeological evidence. No serious alternative was offered to serial endosymbiosis for the origin of the eukaryotic complement of membrane-bound organelles. Discussion concentrated instead on which of these organelles might have come first. Starting with a nucleated host, most preferred first to add mitochondria of probable bacterial origin and then chloroplasts, whose 16S rRNA sequences imply a blue-green algal origin.

Was Galileo 2,000 years too late?

from David W. Hughes

SIDEREUS NUNCIUS was published in Venice in 1610 and revealed to the renaissance world "great, unusual, and remarkable spectacles . . . as observed by Galileo Galilei, Gentleman of Florence, Professor of Mathematics in the University of Padua, with the aid of a spyglass". The title page also announced that the greatest discovery was that of the "four planets swiftly revolving about Jupiter at differing distances and periods, and known to none before the Author recently perceived them and decided that they should be named The Medician Stars". One of Galileo's notebooks records how Jupiter was seen on 7 January 1610 with three companion 'fixed stars', two to the east and one to the west. Galileo concluded, after a week of observation, that "there are three wandering stars around Jupiter, previously invisible to everyone".

Galileo's claim to priority has recently been shown to be unjustified; and it is not a matter of being just pipped at the post, but of some 2,000 years.

Xi Ze-zong, of the Institute for the History of Natural Sciences, Academia Sinica, has recently reported in *Chinese Astronomy and Astrophysics* (5, 242; 1981) that Gan De observed Ganymede in the summer of 365 BC.

Gan De was one of the earliest Chinese astronomers and was an assiduous observer of the heavens, and in particular of the planet Jupiter. He wrote two books, *Treatise on Jupiter and Astrological Prognostications*. Both have been lost but fortunately portions have been preserved in *The Kaiyuan Treatise on Astrology* which was compiled by Qutan Xida between AD 718 and 726. Gan De is quoted as having said:

"In the year of chan yan . . . , Jupiter was in Zi, it rose in the morning and went under in the evening together with the lunar mansions Xunü, Xü and Wei. It was very large and bright. Apparently, there was a small reddish star appended to its side. This is called 'an alliance'."

The reference to the zodiacal division Zi and the lunar mansions reveals that the observation was made in the summer of 365 BC and the use of the term 'in alliance' that the small star is described as a subsidiary of Jupiter. The colour 'chi' is a light red.

Xi Ze-zong puts forward no explanation as to why this reference to a satellite of Jupiter has been overlooked until now.

Which satellite was it? Jupiter has four 'galilean' satellites, Io, Europa, Ganymede and Callisto, and when Jupiter is in opposition, at its closest to Earth, these have magnitudes of 4.9, 5.3, 4.6 and 5.6 respectively (a change of 1.0 in magnitude is equivalent to a factor of 2.5 in brightness and bright objects have low magnitude values). At opposition the maximum

angular separations between the satellites and Jupiter are 2.3, 3.7, 5.9 and 10.1 arc min respectively. Ganymede is thus the brightest moon and can be as much as 5.9 arc min away from Jupiter — one-sixth the apparent diameter of the Moon. Unfortunately, Jupiter then has a magnitude of around -2.6 and is thus 760 times brighter.

Xi Ze-zong used a planetarium to simulate two sources differing in brightness by a factor of 760 and separated by 5.9 arc min. People with good eyesight could detect the fainter object. Xi Ze-zong could, however, have saved himself some trouble by turning to a more recent 'ancient tome'. Admiral William H. Smyth wrote in his *Celestial Cycle*, Vol. 1 (1844) that

"Certain *esprits fort* express surprise that Galileo should have been so gratified by this discovery since they hold that the satellites of Jupiter are often seen with the naked eye and they cite the Apennines and Etna and the West Indies and various other fine-climate places as the spots where such a feat is frequently done."

Smyth concluded, however, that the observer must not only be at a site of fine seeing and excellent atmospheric clarity but must also possess "visual organs of extraordinary power".

Two problems still exist. We cannot be sure whether Gan De saw Ganymede or Callisto. Ganymede is the more probable simply because it is the brighter. The reference to the 'light red' colour is also mystifying. Ganymede is too faint for its colour to be perceived with the naked eye. Even in a telescope it appears white and has a very similar colour to that of Jupiter. □



This the hundredth article written for News and Views by David W. Hughes, a lecturer in physics and astronomy at the University of Sheffield. The first, 'Meteors and Meteorology', was published on 13 June 1970 (Nature 226, 1008).

Molecular recognition and the future of monoclonal antibodies

from David Lane and Hilary Koprowski

A MOST striking revelation to emerge from recent experiments with monoclonal antibodies has been the repeated demonstration of cross-reacting sites on molecules in which conventional serology had failed to detect homology. The most obviously suggestive of these cross-reactions are those which have been found between the large T antigen of simian virus 40 (SV40) and normal proteins of host cells^{1,2}. At least four monoclonal antibodies against the T antigen, which is strongly implicated in the oncogenicity of the virus, can be shown to cross-react very specifically with host proteins, one of which, a ubiquitous protein of molecular weight 68,000, is regulated by cell growth¹.

Other cross-reactions, all equally specific but not all as obviously suggestive, have been detected between various other more or less related molecules. The well characterized rat monoclonal antibody 42-21, for example, binds both Thy-1 antigen and the V α chain of immunoglobulin molecules bearing the TEPC-15 idiotype determinant³. A monoclonal antibody obtained after immunization of mice with herpes virus⁴ binds a glycoprotein specific to the virus and virus-infected cells, but also cross-reacts with neutral glycolipids extracted from normal uninfected cells⁵. Yet other monoclonal antibodies cross-react with two cytoskeletal proteins, tropomyosin and vimentin⁶; and another cross-reacts with Thy-1 antigen and intermediate filaments⁷. Probably yet more apparently capricious cross-reactions have been detected but not reported.

What do they mean? The most exciting possibility is that these antibodies are detecting the conserved recognition sites underlying the specific macromolecular interactions that control most important biological processes. But there are serious difficulties in using monoclonal antibodies as the sole indicator of molecular identity. These difficulties are well known to immunochemists, but they may not be fully appreciated by the growing numbers of workers new to the field of immunochemistry who are now turning increasingly to monoclonal reagents in preference to conventional antisera. Before expanding on the potential of monoclonal antibodies in the exploration of molecular recognition, therefore, we shall reconsider the character of polyclonal antisera and the nature of antibody specificity, and discuss how these bear on

the interpretation of immunological cross-reactions between molecules.

There are two crucial points to be considered, the first of which is the heterogeneity of conventional polyclonal antisera. A conventional antiserum raised in a rabbit against bovine serum albumin, for example, contains a large number of different antibodies directed against different discrete determinants or epitopes on the surface of the bovine albumin molecule. Some of these epitopes will also be present on a related molecule such as human serum albumin, and the two antigens will thus be cross-reactive. Other epitopes, on the other hand, will be present only on one or other albumin molecule. By

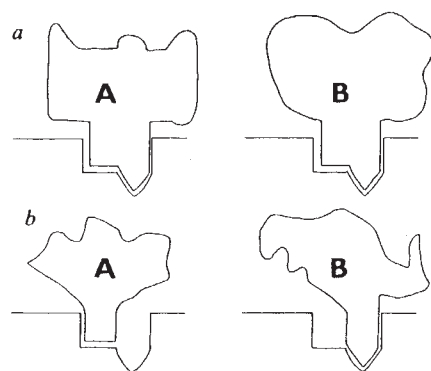


Fig. 1 Two ways a monoclonal antibody could cross-react. *a*, Protein antigens A and B have the same shape over part of their surface. *b*, A and B have different structures but interact with the same antibody molecular.

appropriate absorption experiments polyclonal antisera can be used to distinguish the cross-reactive and the non-cross-reactive epitopes and are thus powerful tools for such purposes as determining phylogenetic relationships between species⁸. Evidently a single monoclonal antibody will never yield this kind of information; and so complex are conventional antibody responses that it would be extremely difficult to mimic them with artificial 'cocktails' of monoclonal antibodies.

But the ease with which homologies between proteins can be detected by conventional antisera depends not only on the intrinsic structure of the antigens but also on the host response to them, which can result in important blind spots in individual sera. For the antibodies produced in a normal immune response are not evenly directed against all parts of the surface structure of the antigen. Rather, most antigens have clearly immunodominant regions which in some cases, for example the influenza haemagglutinin, have been closely mapped⁹. The relative immunodominance of given regions on a protein molecule may

vary between different host animals. For example, different strains of inbred mouse preferentially make antibodies to different areas of the staphylococcal nuclease antigen¹⁰; and certain mutant determinants on human haemoglobin molecules cannot be recognized by rabbit antisera because they cross-react with rabbit haemoglobin, to which the animal is immunologically tolerant¹¹. Such factors will restrict both the ability of individual conventional sera to detect cross-reactions between antigens and, of necessity, the range of monoclonal antibody specificities that can be obtained from a particular cell fusion.

The second important consideration in the interpretation of cross-reactions is directly relevant to the use of monoclonal antibodies, and concerns the nature of the antibody response to a single epitope and the interactions between one pure clonal antibody and that structure. In a conventional antiserum, not only are numerous different antibodies directed against different determinants on the same molecule, but a very large number of different antibodies will react with each single epitope, recognizing it with different degrees of affinity. This extraordinarily diverse response to a single epitope, which may be a very simple artificial determinant such as dinitrophenol, has attracted detailed examination because of its implications for the genetic basis of antibody diversity. A theoretical consideration of the problem led Talmage¹² to propose that one antibody binding site might be polyfunctional and accommodate more than one epitope. He then reasoned that some antibodies would be produced in the immune response to several different antigens and the full range of specificities might be achieved by fewer different antibodies than might at first seem necessary. Thus epitope A would stimulate clones of antibodies with specificity for epitopes A, B, C, D; A, E, F, G; A, H, I, J, and so on. The serum would then bind predominantly epitope A, present on the eliciting antigen. The important implication for monoclonal antibodies is that one would not expect them to show the same degree of A-dominance as the polyclonal serum: they might react equally well with another epitope, for example B, that the whole serum recognizes at such a low titre compared with A as to render it undetectable. This important hypothesis was examined in detail by Richards and colleagues¹³ who obtained direct experimental proof for the existence of polyfunctional binding sites; thus, a myeloma protein (460) was found to bind two structurally unrelated molecules dinitrophenol-lysine and 2-methyl-1, 4-naphthoquinone (menadione) at sites that could be shown to be physically

David Lane is in the Eukaryotic Molecular Genetics Research Group, Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, and Hilary Koprowski is at the Wistar Institute, 36th and Spruce Street, Philadelphia, Pennsylvania 19104.

Bounded brainpower? *from Paul Davies*

Is there a limit, due to fundamental physics, on the rate at which information can be processed in a computer? "Yes" claims Jacob Bekenstein, the co-inventor of the quantum black hole. In a provocative paper published in *Physical Review* a year ago, he claimed that the bread and butter of information transfer — energy and entropy — cannot be varied at will, but are constrained by the bound

$$\frac{\text{entropy}}{\text{energy}} < \frac{k \times \text{size}}{\hbar c} \quad (1)$$

where $\hbar = h/2\pi$, the size of the system is defined in some suitable all-embracing sense, and k , h and c are respectively Boltzmann's constant, Planck's constant and the velocity of light.

Bekenstein went on to assert (*Nature* **292**, 112; 1981) that this bounded ratio of entropy to energy also constrains the rate of information processing to be less than

$$\frac{\pi k \times \text{energy}}{\hbar} \quad (2)$$

Taking into account the need to flush out the heat produced in a hypothetical computer by the information's energy, Bekenstein estimated the upper limit to be about 10^{15} bits per second (still pretty fast).

Now all this has been called into question by David Deutsch of Oxford University's Department of Astrophysics. In *Physical Review Letters* (**48**, 287; 1982), Deutsch challenges the fundamental basis of Bekenstein's analysis — the existence of an entropy to energy bound [equation (1)]. Energy, he points out, is not itself a measurable quantity in non-gravitational physics. Only energy differences are relevant to devices such as computers. Bekenstein's formula (1) might work for systems such as black holes, where gravity plays a part, but it has no business to interfere where

gravity is unimportant. (Energy, having mass, is a source of gravity, so its absolute value can be measured gravitationally.)

Deutsch reworks Bekenstein's calculation for energy differences and finds that there is no upper bound on the entropy to energy difference ratio. He then argues that it is this ratio that is relevant to information processing. The essential point is that not all the energy carried by the information generates heat in the computer, only the energy change due to the encoding procedure. The rest is invisible to everything except the gravitational field. He also throws in, for good measure, the retort that, in any case, it is in principle possible, even in quantum physics, to retrieve information by measurements that leave the information undisturbed, and hence do not produce heat.

So the current state of play is that Bekenstein's bound (1), while not incorporating the newtonian gravitational constant G explicitly, is nevertheless a gravity formula in disguise, to be used only when the system concerned runs on gravity power. In that case, if a computer is too big, its thinking time is limited by the speed of light; too small and it implodes into its own black hole. In this sense, bound (1) may provide a genuine limit, which Deutsch estimates at around 10^{42} bits per second!

Alas, even the gravitational applications of equation (1) have recently been called into question by William Unruh, Robert Wald, Don Page and Stephen Unwin in some lively exchanges (*Physical Review D*, in the press). Evidently the embattled Bekenstein has caught the attention of more than the computer industry.

Paul Davies is Professor in the Department of Theoretical Physics, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

discrete (though the two haptens did compete for the binding site).

In light of these considerations, there are two distinct explanations of the physical basis of monoclonal cross-reactions. One interpretation is illustrated in Fig. 1a, where protein antigens A and B share a small and precise detail of their surface topology. Such a determinant might not normally be detected using conventional antisera in which most of the antibodies will be directed against other A-specific or B-specific structures; only dissection of the antibody response using the monoclonal technique reveals the shared structure. The question then becomes one of just how significant such small homologies are.

In the second explanation (Fig. 1b), the two protein antigens have dissimilar structures but interact with the same antibody molecule. It is notable that the

two different antigens could still compete for binding to the antibody because of steric constraints, even though their reactive epitopes, by analogy with the myeloma 460 example, occupy discrete sites on the molecule. However, the affinity of the antibody for antigens A and B would then be expected to be much lower than in the alternative model illustrated in Fig. 1a. In the second model (Fig. 1b) there is no reason to assume any biological relationship between the cross-reacting antigens; the only factor of interest in the system is the existence of the cross-reactive antibody itself. The production of such antibodies, for instance, could contribute to an autoimmune response induced by a structurally completely unrelated antigen. These two models represent the extreme alternatives and of course it must also be possible for the epitope on antigen A to be

only partly homologous to the epitope on antigen B, so that the antibody binds antigen A with higher affinity. Detection of such cross-reactions, whatever their molecular basis, means that reaction with a monoclonal antibody cannot itself be interpreted as proving molecular identity. For example, the reaction of a monoclonal antibody with two different cell types still requires biochemical verification that an identical molecule is being recognized in each cell. Similarly, such reactions might constrain the use of monoclonal antibodies in radioimmunoassay. Clearly, both of the extreme models have some validity, and examples of both types of cross-reaction will be discovered. Our bias is that the physical basis of the majority of cross-reactions detected so far is nearer to that illustrated in Fig. 1a, because they seem so highly specific. For example, the Pillemer and Weissman antibody shows no detectable binding to Thy-1-negative mutant lymphoma cells or to any IgG except that of the T-15 idiotype, and one of the SV40 T monoclonals has a very high affinity for T yet uniquely recognizes a single low-abundance 68,000 host protein on Western blots of total proliferating cell homogenates. (It does not recognize any protein at all when the extracts are made from the same cells in a quiescent state.)

The question then becomes one of just how functionally significant these small homologies between proteins are. That some are significant is nicely illustrated by the sweet-tasting proteins thaumatin and monellin. Hough and Edwardson¹⁴ have shown that polyclonal antibodies against thaumatin mimic the sweetness receptor, in that other sweet-tasting substances displace ¹²⁵I-labelled thaumatin from the antibody with an efficiency that correlates well with their relative sweetness. Of particular interest is the very effective competition between thaumatin and another very sweet-tasting plant protein, monellin, because the primary sequences of these two proteins show surprisingly little homology. Thaumatin has a single chain of 207 amino acids, and monellin has two chains totalling 94 amino acids; the homology is limited to five identical tripeptides¹⁵. We can infer that the antibody response has been selective for the biologically active region of the molecule, and by concentrating itself in a biologically relevant structure has mimicked a receptor.

A similar example of receptor mimicry by antibodies can be seen in the studies of Sege and Peterson¹⁶ where anti-idiotypic antibodies, raised against antibodies to insulin, were found to possess insulin-like activity themselves in biological tests. Cross-reactive monoclonal antibodies may therefore lead us to discover biologically important relationships between proteins and other macromolecules that could not have been detected by other means. After all, biological macromolecular interactions are all about three-dimensional shape

recognition, which is what antibodies do.

The virus-host interface is an especially interesting area in which one might expect to find such homologies. In the case of the oncogenic RNA viruses, Blair *et al.*¹⁷ have shown the existence in normal mouse DNA of a cellular homologue to Moloney murine sarcoma virus and have demonstrated transformation of murine cells by recombination of the cloned homologous sequences and plasmids containing proviral long terminal repeat sequences of Moloney sarcoma virus. Derivation of *src* sequences of two murine sarcoma viruses, Harvey and Kirsten, from normal genes (*sarc*) of different vertebrate species such as rat, chicken, human and mouse has also been described recently¹⁸. These sequence homologies between retroviral onc genes and cellular genes have not been detected in analogous experiments using the transforming genes of the DNA tumour viruses. Structural homologies between these viruses transforming gene products and host proteins must exist¹⁹, however, and they may now be detected by monoclonal antibodies.

It is our hope that many exciting new correlations between important macromolecules not readily detectable by other means will emerge from these studies and increase our understanding of biological organization. □

1. Lane, D.P. & Hoefler, W.K. *Nature* **288**, 167 (1980).
2. Crawford, L., Leppard, K., Lane, D. & Harlow, E. *J. Virol.* (in the press).
3. Pillemmer, E. & Weissman, I.L. *J. exp. Med.* **153**, 1068 (1981).
4. Frankel, M. *et al.* *Virology* (in the press).
5. Blaszyk, M., Karlsson, K.A., Wroblewska, Z., Frankel, M. & Koprowski, H. (in preparation).
6. Blase, S.H., Matsumura, F. & Lin, J.J.C. *Cold Spring Harb. Symp. quant. Biol.* **49**, Ch. 39a (in the press).
7. Dulbecco, R. *et al.* *Nature* **292**, 772 (1981).
8. Prager, E.M., Wilson, A.C., Lowenstein, J.M. & Sarich, V.M. *Science* **209**, 287 (1980).
9. Lubeck, M.D. & Gerhard, W. *Virology* **113**, 64 (1981).
10. Berzofsky, J.A., Schecter, A.N., Shearer, G.M. & Sachs, D.M. *J. exp. Med.* **145**, 123 (1977).
11. Reichlin, M. *J. molec. Biol.* **64**, 485 (1972).
12. Talmage, D.W. *Science* **129**, 1643 (1959).
13. Richards, F.F., Konigsberg, W.H., Rosenstern, R.W. & Varga, J.M. *Science* **187**, 130 (1975).
14. Hough, C.A.M. & Edwardson, J.A. *Nature* **271**, 381 (1978).
15. Van der Wel, H. *Trends biochem. Sci.* **5**, 122 (1980).
16. Sege, K. & Peterson, P.A. *Proc. natn. Acad. Sci. U.S.A.* **75** (1978).
17. Blair, D.G. *et al.* *Science* **212**, 941 (1981).
18. Ellis, R.W. *et al.* *Nature* **292**, 506 (1981).
19. Lane, D.P. & Crawford, L.V. *Proc. R. Soc. B* **210**, 451 (1980).

Genetic diagnosis of the fetus

from Stuart H. Orkin

THE past decade has seen rapid progress in the development of reliable methods for the identification of genetic disorders in the human fetus. Chromosome abnormalities, such as Down's syndrome, are now routinely detected during mid-trimester by the analysis of fetal cells shed into the amniotic fluid. Structural abnormalities, such as neural tube defects, may be due to many factors but can now often be detected by screening for α -fetoprotein in maternal blood and amniotic fluid. Enzyme deficiencies related to single gene disorders may also be identified from the culture of fetal cells, obtained by the safe and simple aspiration of amniotic fluid in mid-trimester.

Many disorders, however, are not readily detectable in fetal cells from amniotic fluid because expression of the relevant genes is restricted to particular cell types. Conditions affecting either the structure or synthesis of the polypeptide chains of haemoglobin are the best known examples. Sickle cell anaemia, the result of the most common structural abnormality of the human β -globin chain, and β -thalassaemia, the consequence of underproduction of β -globin chains, are disabling genetic disorders of widespread occurrence. Only since the mid-1970's have new methods, the most recent of them employing the techniques of molecular

biology to analyse the genetic material itself, begun to show success in the early diagnosis of these diseases.

The search for suitable diagnostic methods first focused on acquiring fetal blood samples since the β -globin gene is expressed, although at low-levels, during mid-trimester. This approach depended not only on reliable chromatographic methods for assaying the globin products, but also on the development of relatively safe techniques for sampling fetal blood. Direct aspiration of placental veins via a fetoscope has been especially successful although blind aspiration of the placenta has been used as well, with good results¹. In both cases the risk of sampling to the fetus is of the order of five per cent, considerably above that of simple amniotic fluid taps. Such a risk is certainly acceptable in families faced with a one in four chance of having a severely affected child. By virtue of the pioneering work of Nathan, Kan, Kazazian, and their colleagues, the prenatal detection of haemoglobinopathies using fetal blood is now an accepted procedure and has been used in more than 1,800 cases². This, in conjunction with genetic screening and counselling has reduced the birth rate of β -thalassaemics in several regions. As has often been the case in work of this kind, the technical developments have brought additional benefits — it is now possible to detect haemophilia and rare granulocyte dysfunction syndromes *in utero*, and to assess directly fetal cell

chromosomes where amniotic cell analysis may suffer from contamination with maternal cells.

The advantages of determining the presence or absence of a genetic disorder by direct gene analysis, as well as the risks involved in the acquisition of fetal blood, led in 1978 to efforts to apply novel techniques of DNA analysis to prenatal diagnosis³. Some forms of thalassaemia were known from molecular hybridization experiments^{4,5} to be due to structural gene deletions. By using specific DNA probes for globin sequences that could be labelled to high specific activity^{6,7}, bacterial restriction enzymes that cleave DNA at specific nucleotide sequences⁸ and the simple gel method devised by Southern⁹ it was possible to detect such deletions in a DNA sample of 10 μ g or less. This amount of DNA could be obtained from fetal cells either directly on withdrawal of amniotic fluid or after *in vitro* culture. However, although the ability to detect these rare deletion states was a step forwards, for both sickle cell anaemia, which is due to a substitution in codon six of an normal β -chain sequence, and β -thalassaemia, the result of many different single nucleotide substitutions either in the coding or intervening sequences of a normal β -gene¹⁰, a more specific assay was required.

A remarkable finding by Kan and Dozy presented new alternatives. They identified a DNA polymorphism (for the enzyme *HpaI*) in the 3'-flanking region of the β -gene that was quite often linked with the sickle β -gene itself¹¹. By examining the β -gene fragments after *HpaI* digestion in family members, diagnosis of sickle cell anaemia was possible in many, though unfortunately not all, families known by linkage analysis to be at risk¹². Limitation of linkage between the polymorphism and the sickle gene to about 60–70 per cent and its variation in different world populations ultimately restricts the usefulness of this approach. The finding of additional restriction polymorphisms in the β -gene cluster does, however, extend its applicability to sickle cell diagnosis¹³. Similarly, β -thalassaemia may often be diagnosed now by linkage analysis.^{14,15}

Nevertheless, direct detection of the primary gene mutation would be desirable as linkage analysis generally requires careful family studies with several restriction enzymes, an affected family member and often considerable amounts of precious fetal DNA. The power of restriction enzyme mapping is such that a single base change in a restriction enzyme recognition site can be readily detected. Nienhuis first noted that the restriction enzyme *MnII* (recognition site GAGG) would cleave the normal but not the sickle cell codon six¹⁶. However, because *MnII* cleaves β -genes DNA very frequently, experimental use of this enzyme for sickle cell detection has not been successful.

Recently, Geever *et al.*¹⁷ noted that a different restriction enzyme *DdeI* (recog-

nition site CTNAG) would cleave normal, but not sickle, DNA at the sixth codon. Knowing the sequences of the normal β -gene and its immediate flanking region, the sizes of the restriction fragments expected from the normal and sickle cell DNAs could be predicted. Using electrophoresis of DNA fragments in acrylamide rather than the usual agarose, covalent transfer of fragments to a solid support and larger amounts of DNA (20 μ g or more), Geever *et al.* could distinguish the β^S from the normal gene in total DNA. With a similar assay Chang and Kan¹⁸ showed that the approach could be applied to prenatal diagnosis. The detection of the β^S in these studies is an experimental *tour de force* given the relatively small size of the DNA fragments detected (less than 0.4 kb). In view of the technical difficulties of this approach Chang and Kan have advocated its use when the *HpaI* polymorphism is not informative.

Even this achievement may now be superseded — another restriction enzyme has been discovered that cleaves a subset of *DdeI* sites. While it cuts normal β -gene at codon six it does not cleave *DdeI* sites just upstream from the gene. The size of the DNA fragments from normal and sickle DNA should be considerably larger than those obtained using *DdeI* and more suitable for gene mapping. These predictions have recently been confirmed in the present author's laboratory¹⁹. The β^S gene can now be detected directly following digestion with this new enzyme with a sensitivity comparable to detection of the *HpaI* polymorphism in the usual Southern gene mapping analysis. We believe that this constitutes a simple and reliable assay for the sickle mutation in DNA.

The prospects for direct detection of β -thalassaemia mutations are less favourable, however. Some β -thalassaemia mutations can be detected directly by the alteration of the β -gene map by a specific substitution, but most, at least among Mediterraneans, cannot. Linkage analysis remains, however, a useful technique in these situations^{10,14,15}. Other single-gene disorders may, in principle, be identified by linkage analysis using human DNA fragments cloned from various chromosomes, thereby extending this approach to other conditions.

Families seeking prenatal detection of genetic disorders generally choose termination of pregnancy if the fetus is found to be affected with the disease for which it is at risk. The trauma for the family, both physical and psychological, is great, especially as diagnosis is not established until the latter portion of the second trimester. Earlier diagnosis would be very desirable if risks to the fetus or mother were not appreciable. Niazi *et al.*²⁰ and Kazy *et al.*²¹ have reported trophoblast or chorion biopsy by the transcervical route at 6–12 weeks gestation. Such material provides fetal cells, that can be grown *in vitro* if necessary, suitable for

chromosome, enzyme and DNA analyses²². Although early study has shown no complications associated with the procedure, more investigation will be needed before it gains wide acceptance. Nevertheless, the prospect of obtaining fetal cells in the first trimester is exciting indeed. It will obviate the "race to the wire" that sometimes prevails in mid-trimester diagnoses and may ultimately alleviate considerable psychological and physical morbidity. The availability of such procedures will provide further incentive to find methods for DNA analysis, either direct or by linkage, for many genetic diseases.

Finally, we may wonder whether totally noninvasive approaches to diagnosis of disease in the fetus might not be possible. Fetal cells normally pass in small numbers into the maternal circulation. Specific enrichment of fetal cells by cell sorting, perhaps using recently developed monoclonal cell surface-specific antibodies²³, may eventually allow fetal diseases to be detected in cells obtained from maternal blood. If this is likely to be useful in situations where DNA is analysed, it may be necessary to design novel gene

probes that extend the sensitivity of present gene mapping methods. At the current pace of progress in this field, we can anticipate that we will soon be surprised by as yet unforeseen developments. □

1. Benzie, R.J. in *Laboratory Investigation of Fetal Disease* (ed. Barson, A.J.) 320 (Wright, Bristol, 1981).
2. Alter, B.P. *Lancet* ii, 1152 (1981).
3. Orkin, S.H. *et al. New Engl. J. Med.* **299**, 166 (1978).
4. Ottolenghi, S. *et al. Nature* **251**, 389 (1974).
5. Taylor, J.M. *et al. Nature* **251**, 392 (1974).
6. Wilson, J.T. *et al. Nucleic Acids Res.* **5**, 563 (1978).
7. Little, P.F.R. *et al. Nature* **273**, 640 (1978).
8. Roberts, R. *Meth. Enzym.* **65**, 1 (1980).
9. Southern, E.M. *J. molec. Biol.* **98**, 53 (1975).
10. Orkin, S.H. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 2400 (1979); Flavell, R.A. *et al. Nucleic Acids Res.* **6**, 2749; 1979; Orkin, S.H. *et al. Nature* (in the press).
11. Kan, Y.W. & Dozy, A.M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5631 (1978).
12. Kan, Y.W. & Dozy, A.M. *Lancet* ii, 910 (1978).
13. Phillips, J.A. III *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2853 (1980).
14. Little, P.F.R. *et al. Nature* **285**, 144 (1980).
15. Kazanian, H.H. Jr *et al. Blood* **56**, 926 (1980).
16. Nienhuis, A.W. *New Engl. J. Med.* **299**, 195 (1978).
17. Geever, R.F. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 5081 (1981).
18. Chang, J.C. & Kan, Y.W. *Lancet* ii, 1127 (1981).
19. Orkin, S.H., Little, P.F.R., Kazanian, H.H. Jr & Boehm, C.D. (in preparation).
20. Niazi, M., Coleman, D.V. & Loeffler, F.R. *Br. J. Obstet. Gynecol.* (in the press).
21. Kazy, Z., Rozovsky, I.S. & Bakharev, V.A. (in the press).
22. Williamson, R. *et al. Lancet* ii, 1125 (1981).
23. Lipinski, M., Parks, D.R., Rouse, R.V. & Herzenberg, L.A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5147 (1981).

Trace elements and ocean chemistry

from Michael L. Bender

THE article in this issue of *Nature* by Elderfield and Greaves (p.214) reflects two important trends in the study of the marine chemistry of the trace metals. First, there is increasing interest in the detailed chemical behavior of trace elements in the oceans — the complexing of metals with anions and organic ligands, particle adsorption and the occurrence and the implications of trace metal redox reactions. Second, we can see the trend towards application of four of the sacred instruments of hard rock geochemistry — lead, strontium and neodymium isotopes and rare earth element distributions — to problems in marine geochemistry.

The earliest successful research on marine trace metals, begun in the early 1960s, concentrated on elements whose distribution in seawater could be accurately measured by the radiochemical study of uranium and its daughter Ra²²⁶. Subsequently, the approach has been extended to include the less abundant radioisotopes of thorium, radium, protactinium, polonium and lead. In the mid-1970's, a number of laboratories began to study the transition metals manganese, iron, nickel, copper, cadmium, zinc and chromium, which are of great interest because they are both essential and toxic to organisms, and

because they are enriched in both hydrothermal and normal deep sea sediments. During the last few years, considerable attention has been paid to the metalloids tin, selenium and germanium, and the methylated compounds which they form. Finally, extensive effort has gone into determining the oceanic distributions of the global pollutants, lead and mercury.

Such studies have given us a broad understanding of marine trace element geochemistry. A number of elements, including thorium, protactinium, polonium, lead, iron and manganese, are highly reactive and are scavenged from seawater into sediments within a few decades. Others, including copper, nickel, cadmium, zinc, selenium, and cobalt, are rapidly removed from surface water by sinking tissue, but redissolved when the tissue is oxidized in the deep sea. They remain in the oceans for 10³–10⁵ years before their final removal into the sediments. Selenium and chromium are present in the dissolved form in two oxidation states, while manganese and iron are rapidly scavenged once the reduced species are oxidized. The ocean's sources of trace elements — input from rivers, submarine hydrothermal processes (at least for manganese), and dissolution from aerosols (for manganese, lead, and copper among other elements) — are fairly well known. The removal mechanisms are less well understood, although it is clear that burial

Michael L. Bender is in the Graduate School of Oceanography University of Rhode Island Kingston, Rhode Island 02881.

of trace metals carried to the sediments by falling tissue is the dominant process.

Most of our understanding of the marine cycles of trace metals is based on concentration variations in seawater. Such data cannot, however, give us a profound understanding of the physical chemistry governing their behavior. To meet this problem, some groups have studied the inorganic ion pairing reactions of the trace metals, metal-organic interactions, and the reactions between trace metals and important inorganic and organic surface-active phases. Other groups are studying trace metal behavior by radiolabelling trace metals in enclosed ecosystems, and determining their uptake rates and solid phase distributions.

The rare earth elements (REE) have a unique role to play in extending our understanding of marine trace metal geochemistry. Their ionic radii, and hence ion pairing and surface adsorption stability constants, vary in a predictable way with atomic number. As a consequence the importance of complex formation and adsorption in the marine cycles may be inferred directly from the subtle interelement differences in REE distributions. While all REE exist in the +3 oxidation state, Ce and Eu also exist in the +4 and +2 oxidation states, respectively. By comparing their distributions with those of adjacent rare earths, it is possible to judge how their behavior depends on oxidation state. The relative abundances of rare earth elements are distinctly different in various continental source materials, seawater, and sediments and the abundance patterns can serve as 'fingerprints' that identify sources of the elements in seawater, sediments, and manganese nodules. The isotope composition of one of the REE's, neodymium, itself serves as a fingerprint, since it varies in different source rocks as a result of Sm radiodecay to Nd¹⁴³. Finally, the power of all these tools for studying REE chemistry is enhanced, as Elderfield and Greaves show, by two happy coincidences. First, REE are affected by nearly all of the important processes influencing marine trace metal behavior: atmospheric input, adsorption and scavenging onto falling particles, coprecipitation, ion pairing, and redox reactions. Second, the scavenging residence time for REE — apparently about 400y — is shorter than the oceanic mixing time (1,600y), so that REE and Nd isotopes always reflect elemental sources. If the residence time were longer than the mixing time, the fingerprinting would be less distinct and the REE studies would be less useful.

The second trend reflected in Elderfield and Greaves article can be seen in a number of recent papers. Pb isotopes are now being used to trace the anthropogenic lead distribution, while Sr isotopes are being used to study low and high temperature seawater-basalt exchange as well as

carbonate diagenesis. Elderfield *et al.* (*Geochim. cosmochim. Acta* **45**, 1231; 1981) used REE distributions to characterize the source of these elements in North Pacific nodules and sediments. Goldstein and O'Nions (*Nature* **292**, 3241; 1981) have used the strontium and neodymium isotope compositions to show that, in manganese nodules, those elements have been precipitated from seawater, while Piegras and Wasserburg (*Earth planet. Sci. Lett.* **50**, 128; 1981) have shown that the isotope composition of dissolved neodymium is different in the Atlantic and Pacific oceans (indicating a residence time less than the mixing time).

Rare earth concentrations in seawater were first measured successfully in a few samples during the sixties by Goldberg *et al.* (*J. geophys. Res.* **68**, 4209; 1963) and Högdahl *et al.* (*Adv. Chem. Ser.* **73**, 308; 1968.) Since then only a small amount of additional data has been published. The

paper by Elderfield and Greaves represents a major advance in the accuracy of measurement, and through the study of a complete water column profile allows an appreciation of the roles of particles in removing REE from the shallow waters of the ocean and remineralization redissolving them at depth.

The data allows Elderfield and Greaves to outline the main aspects of both the long term mass balance of REE in the oceans, and the pathways involved in the internal cycling of REE. The context is thus set for more detailed studies of the oceanic distributions of the REE. The results promise to give us a profound understanding of the mass balance, internal cycling, and physical chemistry of REE's in seawater, and thereby greatly extend our appreciation of how the fundamental chemical properties of trace metals influence their behavior in the marine environment. □



100 years ago

THE CHANNEL TUNNEL

The two schemes for a tunnel beneath the Channel, on the comparative merits of which a Parliamentary Committee will probably take evidence in the course of the year, are based, like those which have preceded them, chiefly on geological considerations. The Weald of Kent and the Bas-Boulonnais, once in all probability geographically continuous, still constitute a single geological area.

The two schemes which are offered for tunnelling through the Chalk may be briefly stated as follows:—

1. The Channel Tunnel Company, with which Sir John Hawkshaw is connected, propose a tunnel starting from Biggin Street, Dover, with a gradient of 1 in 80, passing under the north spur of Dover Castle Hill, and thence continuing to a point on the shore known as the Fan Hole at a distance of 2 miles, 4

furlongs, 2150 chains from Biggin Street, and at a depth of 115 feet below high water ordinary spring tides. From this point it will run across the straits to join the French tunnel, which commences near Sangatte, and as Sir J. Hawkshaw has always advocated a straight line of tunnel, we presume that such will be the case here.

2. The Submarine Continental Railway Company, with which Sir E. Watkin is associated, propose a tunnel connected with the South-Eastern Railway, about two furlongs west of Folkestone entrance of the Abbotcliffe Tunnel by a tunnel descending at a gradient of 1 in 52 to the bottom of the No.2 Shaft, near the west end of the Shakespeare Tunnel, at a depth of 126 feet below high water ordinary spring tides. From this point the tunnel will continue for about a mile towards the head of the pier in a direction slightly diverging from the shore, and finally curving round, will fall into the line of the French tunnel near Sangatte.

So far as our information goes at present it seems that the most serious obstacle will be water, and it is therefore on their relative liability to flooding that the proposed tunnels must be judged.

From *Nature* **25**, 464-465; March 16, 1882.

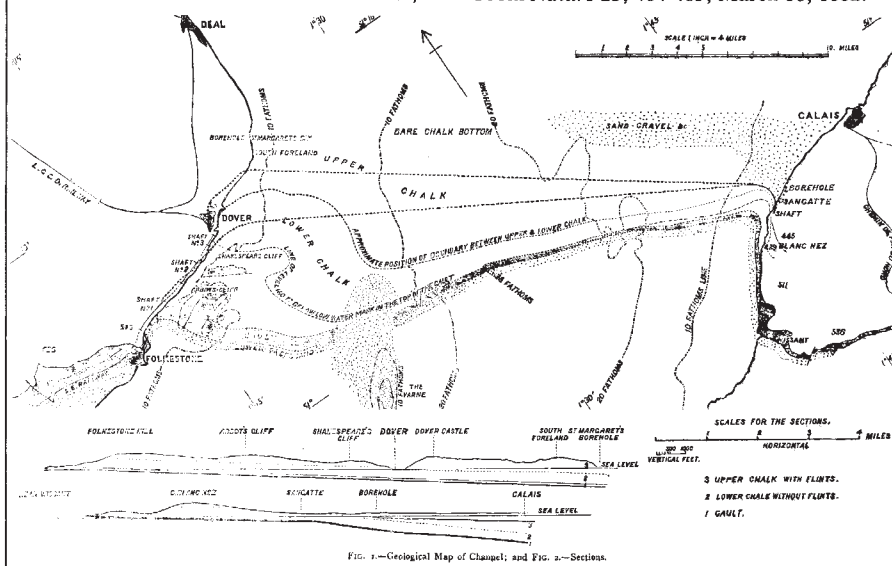


FIG. 1.—Geological Map of Channel; and FIG. 2.—Sections.

REVIEW ARTICLE

The map sense of pigeons

James L. Gould

Department of Biology, Princeton University, Princeton, New Jersey 08544, USA

Homing pigeons are able to return to their lofts after displacement to unfamiliar sites hundreds of kilometres away. A map sense enables them to determine their location, and a hierarchy of compass senses orient their flight. Recent experiments implicate magnetic and olfactory cues as the basis of the mysterious map sense.

HOMING pigeons can be taken from their lofts and transported hundreds of kilometres in covered cages to unfamiliar sites and yet, when released, be able to choose fairly accurate homeward bearings within a minute and fly home. From direct tracking we know that a homing pigeon typically circles near the release site once or twice, edging in the general direction of home, and then straightens out and flies off at nearly the correct bearing¹ (Fig. 1a). At most sites the usual vanishing bearing—the direction in which the bird is lost to view or, if it is carrying a radio transmitter, in which its signal fades out—is reliably left or right of the true home bearing. This consistent deviation is known as ‘release-site bias’^{2,3} since it varies from site to site (see Fig. 1b) or as ‘preferred direction’^{4,5} since it seems to impose a fairly consistent directional tendency onto the homeward direction of birds from any one loft (NNW in Fig. 1b). However, the exact route home is quite variable, and even the same bird released time and time again from the same site never seems to fly the same route twice. Once near home pigeons rely on visual landmarks to reach the loft⁶. Detailed vision is not necessary until the very last stages of the homing process, however, since birds wearing frosted lenses can get to within a few kilometres of home⁷ (Fig. 1c). Moreover, pigeons can even be trained (though with great difficulty and high attrition) to home at night⁸. Pigeons are not unique in their ability to locate precise targets from great distances—animals as diverse as monarch butterflies and sea turtles perform similar feats as part of their normal migratory repertoire—but the willingness of the pigeon to home on demand, day after day, has made it the favoured vertebrate for navigational research.

Aside from reading the minds of the experimenters (a possibility which has not escaped investigation), there are only three basic explanations for the remarkable ability of pigeons to home⁹: (1) the birds might keep track of their outward displacement (the system of many short-range species such as honey bees); (2) they might follow some home-centred gradient to their goal; or (3) they might have a true map sense which would permit them to ‘place’ themselves with respect to home on some internalized coordinate system.

For pigeons (and probably most other wide-ranging animals) the first two alternatives seem unlikely. If pigeons had to measure their displacement (presumably by keeping track of the direction and degree of acceleration and deceleration of the various turns, and timing the individual legs of the journey), simply transporting them in the dark or with constant rotations ought to impair or eliminate their ability to orient. This treatment, however, has no effect^{10–12}, nor does transport under complete anaesthesia^{10,13,14}, or bisection or surgical removal of the vestibular apparatus^{15–19}. Another model for an inertial system might involve a magnetic compass to measure the direction of each leg of a journey, and a time sense to estimate the length of each part. Birds transported to the release site wearing magnets or otherwise subjected to an artificial magnetic field, however, are often not affected (refs 8, 11, 12; but see below). Unfortunately, no one has yet performed the crucial experiment of transporting pigeons in total darkness anesthetized, rotating, and with the field reversed all at the same time.

The gradient alternative is also entirely lacking in experimental support. Such a strategy, based as it is on a ‘getting warmer’ system of movement, requires no compass sense. Yet pigeons have been shown to have compasses and to depend on them for homing; as we shall see, birds may be made to fly directly away from home by such relatively simple manipulations of their compass sense as clock shifts.

The compass sense

The idea that pigeons must have both map and compass was first put forward by Kramer²⁰. As he pointed out, an animal transported to an unfamiliar site probably needs both: a map sense to determine the direction of displacement from home, and a compass to find that direction. (In theory, of course, a pigeon could home by means of successive map readings without any recourse to a compass; but the clock shifts described below, as well as the ability of the birds to select a reasonably accurate homeward bearing within 100–500 m of the release site, render this possibility extremely unlikely.) Kramer demonstrated that the usual daytime compass of birds is the Sun. He trained starlings²¹ (and subsequently a variety of other birds including pigeons²²) which had a view of the natural Sun to seek food in a specific compass direction, and they readily learned to perform all day long. Apparently the birds could compensate for the Sun’s westward movement in the sky. When Kramer deflected the Sun’s image with mirrors, the birds were reoriented in the concomitant direction. Finally, when Hoffmann shifted the circadian rhythm of his birds—so that subjective dawn was at noon, say—the birds misread the Sun’s azimuth accordingly²³. Extensive work with clock-shifted pigeons in the field has confirmed that their preferred daytime compass is solar: a bird shifted by 6 hours will depart from a release site roughly 90° off the homeward bearing on a sunny day^{24–26}.

Keeton found that clock-shifted pigeons orient and home normally under overcast skies²⁵; this raises the possibility that rather than using celestial cues perceived through the clouds, they must utilize some back-up compass. Merkel and his colleagues had already shown an exceedingly weak but persistent ability in caged migrants to orient appropriately to the Earth’s magnetic field^{27,28}. Keeton discovered that pigeons carrying magnets on their upper backs could orient normally on sunny days, but usually not under overcast skies²⁹. Walcott and Green³⁰, using pigeons equipped with miniature Helmholtz coils on their heads, were able to extend Keeton’s results, lending support to the Wiltschkos’ suggestion (based on caged migratory robins³¹) that ‘north’ for pigeons is the direction of the dip lines into the Earth independent of actual magnetic polarity. Given the crucial importance of this ‘dip-angle’ hypothesis it is remarkable that the Wiltschkos’ experiment has not been repeated. Walcott’s basic results have been replicated by Visalberghi and Alleva³².

More recently, the Wiltschkos have followed up Keeton’s finding that young birds wearing magnets on their first flights cannot orient even with a full view of the Sun²⁹. They have discovered that the Sun compass system must first be calibrated

against the magnetic compass³³. Hence, pigeons kept under permanent clock shift since birth are able to orient and home perfectly well, but are misoriented after being reset to normal time³⁴. The emerging picture of the pigeon compass, then, is that the magnetic sense is innate—based perhaps on the deposits of magnetite recently discovered in pigeons^{35,36} and other navigating species—and is used to calibrate a time-compensated Sun compass³⁴. Having served as an essential reference, the magnetic compass sense seems to be largely or completely ignored whenever solar information is available. A very similar pattern of calibration has now been documented in the ontogeny of orientation in nocturnal migrants^{37–39}.

The map sense

In sharp contrast to the steady progress in our understanding of the nuts and bolts of the pigeon compasses, the map sense of

these birds has proven exceedingly difficult to elucidate. The first serious theory was put forward by Yeagley^{40,41}. He imagined that pigeons determine their location by means of a bicoordinate grid of magnetic field intensity and Coriolis force. Since the north magnetic pole is about 2,300 km from the north geographic pole (the axis of rotation), the isolines of vertical magnetic field strength and Coriolis forces intersect at about a 30° angle in the northeastern United States. Yeagley's hypothesis has been discounted on both experimental and theoretical grounds⁴².

The second major effort to explain the pigeon's map sense was by Matthews^{43,44}. He proposed that birds could use the Sun's elevation and arc (which they would extrapolate through the sky on the basis of its motion over a brief interval) in combination with their own internal time sense. A pigeon taken south-west, for example, will find the Sun's position unusually high and east (which is to say early) by an amount corresponding to its displacement from the loft. As a compass, the pigeon could use the extrapolated highest point in the arc as a measure of south. A variety of simple observations have made this 'Sun arc' hypothesis seem very unlikely. The two most compelling are the ability of pigeons to home under overcast²⁵ and the birds' consistent interpretation of clock shifts as compass rotations and never as map displacements^{24–26}. Hence a pigeon clock-shifted 6 hours early and released at local dawn when its internal clock reads 'noon' ought, according to Matthews, to presume it has been taken a quarter of the way around the world to the west. Such an animal ought to fly east. Instead, it will inevitably depart 90° left of home—west if it has been taken south. Despite its lack of supporting evidence, however, Matthews' hypothesis with its explicit, testable predictions has been the single most important element contributing to our present understanding of both maps and compasses.

At present there are two main map hypotheses, one based on olfaction, the other on magnetism. Both schools of thought offer intriguing and apparently reasonable experimental evidence, and yet neither can, as yet, satisfactorily explain the behaviour of pigeons.

Olfactory maps

The olfaction hypothesis grew out of a curious set of 'palisade' experiments begun by Kramer⁴⁵ and continued by Wallraff^{46–51}. The usual protocol for raising homing pigeons for research is to allow the birds a free view of their surroundings from the loft (presumed to be important for calibration of the celestial compass) combined with daily exercise flights in the immediate vicinity of the loft to build endurance. Finally the birds are taken ever further from the loft for a series of practice releases before testing at longer distances. The practice releases help weed out poor homers. In the palisade experiments, however, the pigeons are never allowed exercise flights or practice releases. As a result, the birds' first release in palisade experiments is usually also their last; such a small proportion is ever found again that data on the directions in which they are found should be discarded. Only the vanishing bearing results can be considered very reliable.

The essential result of the original palisade experiments was that blocking a view of the horizon and the lowest 3° of the sky with wood or even clear glass resulted in birds less well oriented than those permitted an unobstructed view of the Sun, sky and horizon. Blocking the view of the sky with a roof down to just a few degrees above the horizon (and obscuring even that near dawn and dusk to prevent any direct view of the Sun) gave intermediate results (Fig. 2). The conclusion Wallraff drew from these first experiments⁴⁶ was that there must be some non-visual "substrate of information spreading in the horizontal plane which is important for homing".

From data derived from later experiments, Wallraff concluded that birds in cages surrounded by opaque plastic louveres which were supposed to block visual cues but permit air to flow through freely were well oriented⁴⁷, but the original data (with the sporadic recoveries omitted) suggest just the opposite

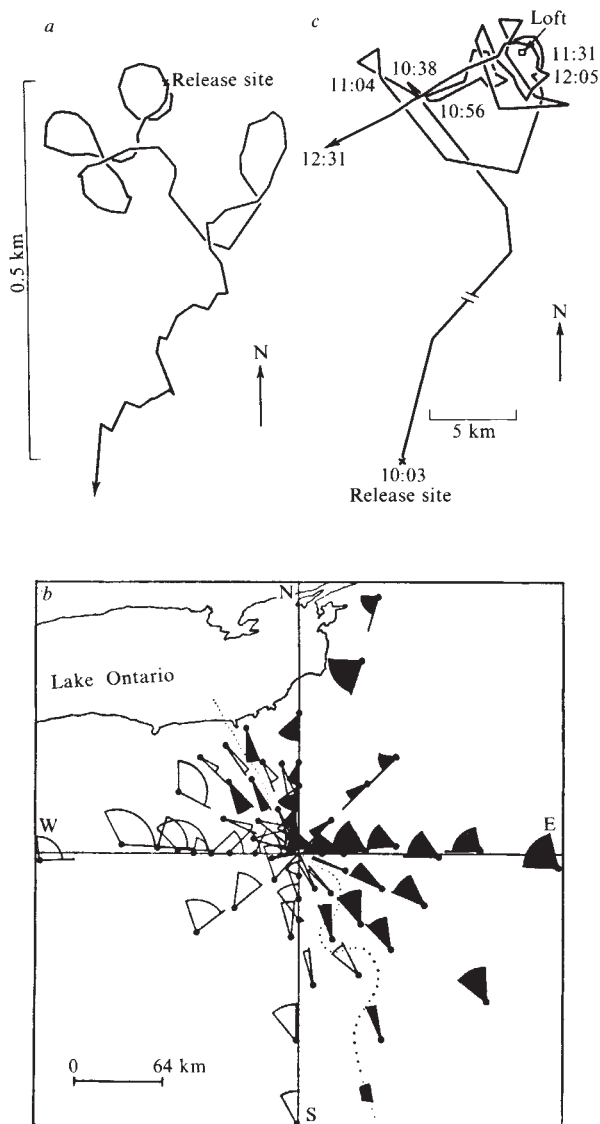


Fig. 1 Behaviour of pigeons while homing. *a*, After release, pigeons typically circle briefly and then fly off in the general direction of home. *b*, Release site biases near Ithaca (centre of grid). The mean bearings of pigeons deviate reliably from the homeward direction at almost all locations. There is a strong tendency for the biases of birds released near Ithaca to be clockwise (black) to the east and counter-clockwise (white) to the west of a line running roughly NNW through Ithaca. *c*, Once near home, pigeons generally use visual landmarks. In this case, however, the bird was wearing frosted contact lenses and so was unable to locate the loft more precisely than a few kilometres. This probably represents the limit of resolution in the pigeon map system. Data from (a) Elsner¹, (b) Windsor², (c) Schmidt-Koenig and Walcott⁷.

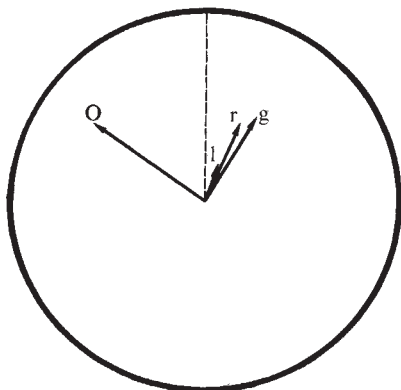


Fig. 2 Results of palisade experiments. By convention, the dashed line is the homeward direction, while the vectors describe the mean angle and degree of scatter among the birds. A vector reaching the circle would represent perfect orientation whereas a vector with no length would correspond to perfectly random data. O, pigeons raised in an open wire aviary (fairly well oriented); r, birds reared in a roofed palisade (well oriented but more scattered); g, pigeons kept in a glass palisade (fairly well oriented but scattered); l, birds raised in a louvred palisade (poorly oriented and highly scattered). Data from Wallraff^{46,47}.

(Fig. 2). Nevertheless those data, taken together with a recent series of elaborate louvred loft experiments⁴⁸ have been interpreted by Wallraff as implicating an 'atmospheric factor'. In these more recent tests birds in (1) a completely exposed caged loft (the controls) were compared to those raised in lofts surrounded to 10° above the horizon by (2) clear glass, (3) opaque plastic louvres, (4) clear glass on two opposite sides with the other two sides open or (5) wood on two sides and opaque plastic louvres on the other two. Wallraff's conclusion was that birds in lofts with a free flow of air (the controls and loft types 3, 4 and 5) orient better than birds in the loft surrounded by glass. The data in this highly regarded set of experiments are confusing at best, and are made all the more so by the lumping together of bearings at 20 seconds, 40 seconds and at vanishing time after release, and the exclusion of birds judged to be flying too low or not far enough away. Most other researchers consider only vanishing bearings to be reliable. Furthermore Wallraff lumps data from four different release sites with extraordinary release-site biases (based on control releases) ranging from 84° right of home to 149° left. The only procedure by which the confounding effect of these biases can be sorted out is to refer the experimental data to the (very limited set of) control bearings on a site-by-site basis (Table 1). On this basis it seems obvious that both the glass and louvred lofts increase the scatter of departure bearings but have little effect on net homeward orientation. The hybrid lofts, on the other hand, severely disorient their occupants. There is no pattern to the effect even when the direction of the louvred axis is compared to releases along or orthogonal to that axis (see below). In sum, then, Wallraff has shown an extraordinary and mysterious synergism between louvres and wood as well as between glass and open sides. These effects are generally interpreted as disrupting the pigeon map sense. As pointed out below, however, a compass effect is at least as likely to be the explanation, and these releases need to be repeated under overcast skies to see if it is the Sun compass that is being affected.

Despite the many inconsistencies in the various palisade experiments, Papi has followed up the seemingly obvious interpretation that the hypothetical 'horizontal, atmospheric factor' might be odour. Papi and his colleagues believe that birds come to associate the odours borne on the wind with the direction in which the wind is blowing, and so slowly build up an olfactory map of their surroundings⁴⁹. When transported to a release site, then, they have only to sniff the air en route and/or at the site to know the direction of home. Hence the louvred palisades ought not to affect orientation while glass ones might.

How odour could be used by naive pigeons at long distances (hundreds or thousands of kilometres), however, remains a total

mystery^{8,19}. Nevertheless, Papi has conducted a long and ingenious series of highly varied experiments whose results readily lend themselves to this olfaction hypothesis: (1) pigeons whose nostrils have been plugged⁵⁰ or whose olfactory nerves have been cut^{51,52} are poorly oriented at release and home slowly; (2) pigeons taken by different circuitous routes to release sites depart in directions slightly favouring the general direction of the particular detour to which they were subjected^{51,53}; (3) pigeons raised with two different artificial odours added to the wind, one from each of two different directions, and then released with one of the odours applied to their nostrils, are deflected from home as though the direction from which they were released corresponded to the direction from which the odour was carried to them in the loft⁵⁴; (4) the release bearings of pigeons raised in 'deflector' lofts (Fig. 3) whose clear glass or plastic deflectors rotate the incoming wind either clockwise or counter-clockwise are similarly rotated^{55,56}; (5) applying a strong odour to the nostrils to block the perception of local odours disorients the pigeons at release⁵⁷; and (6) the vanishing bearings of pigeons raised in lofts rigged with fans to reverse the natural direction of air flow are either roughly reversed or highly scattered⁵⁸.

Problems with the olfactory hypothesis

Taken together, Papi's data seem almost overwhelming. Nevertheless, most of the experiments admit of simpler, nonolfactory explanations, or cannot be repeated, or both. One of the problems with the hypothesis comes from the failure of Schmidt-Koenig and Phillips⁵⁹ to detect any ability in pigeons to distinguish natural air (presumably laden with olfactory map information) from pure, filtered air. The technique they used, cardiac conditioning, is extraordinarily sensitive, and has been successfully used in the past to reveal the sensitivity of pigeons to UV light⁶⁰, infrasonic sounds⁶¹, minute differences in air pressure⁶², and polarized light⁶³, though oddly enough it has never succeeded in picking up any sensitivity to magnetic fields (see, for example, ref. 62). Keeton and his colleagues⁶⁴ have repeated the nerve sectioning experiments but found little or no effects on initial orientation. Indeed, when Benevenuti, one of Papi's colleagues, attempted to repeat the experiment, no effect on initial orientation was evident, though homing speed was slower⁶⁵.

More recently Wallraff⁶⁶ also found that nerve-sectioned birds were less well oriented and slower to home than untreated birds. However, in none of these many experiments by the various groups was homeward orientation abolished. In Wallraff's extensive study, for example, the average mean vector length was an impressive 0.8 for controls, but still a respectable 0.4 for experimentals. Moreover, the orientation of nerve-sectioned birds differed from the controls by only 38° on

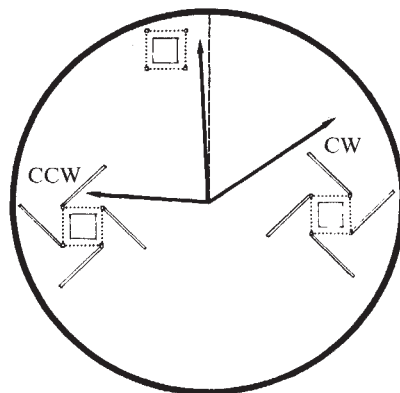


Fig. 3 Design and effect of deflector lofts. The departure bearings of pigeons kept for at least a few weeks in deflector lofts are rotated from those of controls in the same direction as the wind entering the loft is deflected. CW, clockwise loft; CCW, counter-clockwise loft. Data from Baldaccini *et al.*⁵⁵.

average, while groups of controls at the same sites in two different series of releases differed from each other by 41° . In short, these effects are clearly quantitative rather than qualitative, and probably have nothing to do with destroying the map system *per se*. Indeed, tracking reveals that the few birds in these experiments usually land and probably spend extended periods grounded⁶⁴, in contrast to the disoriented birds in other experiments who continue gamely flying. It seems likely that the behaviour of both the nostril-plugged and nerve-sectioned birds results from a generalized, distracting trauma. Indeed, when Keeton used nasal tubes to bypass the olfactory chamber and minimize trauma, no disorientation was evident⁶⁷. Moreover, both Schmidt-Koenig and Phillips⁵⁹ as well as Kiepenheuer⁶⁸ have shown that when the olfactory epithelium of a pigeon is sprayed with the potent local anaesthetic xylocaine, initial orientation is normal. Experiments with cardiac conditioning demonstrate that xylocaine-treated birds are unable to respond even to strong artificial odours for about 90 min after treatment⁵⁹.

Papi's detour experiments have also proven difficult to replicate. Keeton^{64,69} found no real effect in almost all cases while Schmidt-Koenig and his colleagues drew a blank in 11 of 12 tries⁷⁰. At the rare sites showing an effect, however, the deflection was relatively consistent. Of course the olfaction hypothesis is by no means the only explanation for a detour effect. Similarly, neither Keeton's group^{64,71} nor Schmidt-Koenig's⁷⁰ could find that applying masking odours to the beaks and nostrils of pigeons had any predictable effect.

Of all the non-surgical manipulations reported by Papi's group, only the deflector loft has worked consistently elsewhere. Both Kiepenheuer⁷² and Keeton's group^{73,74} have obtained consistent rotations of departure bearings, but there seems no reason to suppose that these dramatic effects have anything to do with odour. Wallraff, for example, has pointed out that the results from his many louvered palisade experiments contradict the olfaction hypothesis. In the heroic series of experiments summarized in Table 1, releases from along the axis of the louveres or open ends of loft types 4 and 5 ought to have yielded far better orientation than releases from along the orthogonal axis. In fact, there was no difference: the mean vector length and angular error for on-axis releases was 0.47 and 99° versus 0.45 and 99° off-axis⁴⁸. Similarly, Kiepenheuer found the deflector lofts worked just as well on birds with plugged nostrils or with xylocaine treatment⁶⁸, implying that the deflection is not a result of rotated windborne odours.

The greatest difficulties with both the palisade and deflector loft experiments arise from the uncontrolled effects that the various clear and opaque plastic and glass elements certainly have on the transmitted and/or reflected UV and polarized light which pigeons perceive. The UV-polarized light of the blue sky is well known to be of supreme importance in the orientation of a variety of animals, most notably the honey bee, and all of the various plastic and glass elements are opaque to UV and greatly alter the degree and direction of polarization. Even more problematical are the reflections each would inevitably generate. Following up on these worries and the suspicious observation that the deflector effect is very slight or even absent under

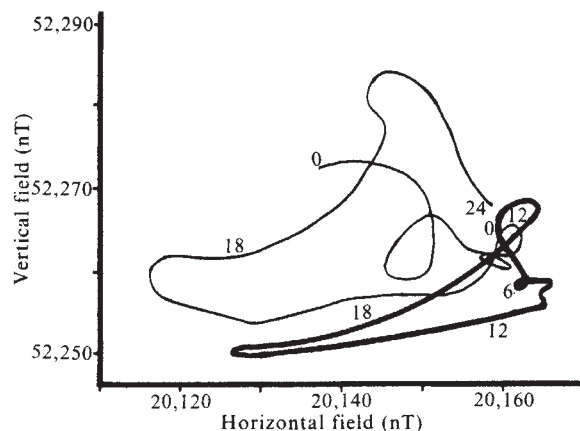


Fig. 4 Daily patterns of magnetic field intensity. On normal days (heavy line) the pattern is relatively regular. (The numbers by the curves are hours GMT.) On 'stormy' days, however, the pattern is irregular (thin line). These curves are based on hourly means recorded by the US Coast and Geodetic Survey and Fredricksburg, Virginia in September 1973. The hourly means suppress the frequent but brief 100 nT pulses which accompanied this storm.

overcast skies, two of Keeton's colleagues, Waldvogel and Phillips, designed a clever deflector loft which rotates wind one direction while it reflects incoming visual cues the other way. The release bearings of the pigeons raised in these lofts are strongly rotated in accordance with the visual cues, demonstrating that the rotation of visual cues is more important in creating the phenomenon⁷⁵. Hence, the deflector loft effect is probably a recalibration of the pigeon compass rather than of the map sense.

A more detailed look at these new data⁷⁵ makes it clear that, as Kramer and Wallraff had already shown, no direct view of the Sun (or its reflection) is necessary for recalibration. This observation points to the probable use of polarized light, perhaps along the lines described by Brines⁷⁶. The most likely calibration component would be the prominent band of horizontal polarization which defines the solar and antisolar meridian, and which Brines and Gould have found to play an important role in bee navigation⁷⁷⁻⁷⁹. Similarly, Able has found that when nocturnal migrants calibrate their compasses at dusk for their upcoming flight, altering the skylight polarization above them rotates the birds' compasses⁸⁰. It seems likely, then, that polarization patterns, so long known to be crucial to the orientation of invertebrates, amphibians, and fish, play a hitherto unappreciated role in birds.

The olfactory hypothesis, though initially very exciting, has lost some of its lustre as a result of these many problems. It may yet turn out to be true, and two experiments—training pigeons to strong, artificial odours in the loft⁵⁴, and reversing the natural wind direction in the loft⁵⁸—still stand. Even here, however, doubts intrude: (1) no one has yet repeated either; (2) Papi found that homing speeds in the strong-odour experiment were unaffected⁵⁴; (3) even Papi's group finds that rearing birds with no olfactory experience whatsoever does not affect initial orientation⁸¹; and finally (4) the maze of 'clear' plastic and glass used to construct these elaborate lofts must introduce just the sort of visual problems which may well account for the deflector effect. Of course there remains the possibility that Italian homing pigeons have a map system altogether different from that of German and American birds; and strain-specific differences in homing performance and initial orientation are well known (W. T. Keeton, personal communication). However, neither Kiepenheuer⁷² nor Wallraff^{11,12} have been able to find any major or consistent differences in orientation between German and Italian pigeons.

Magnetic maps

The major alternative to the olfaction hypothesis at the moment, put forward independently by Walcott⁸², Gould⁹, and Moore⁸³,

Table 1 Comparison of data on differing release sites

Type of loft	Mean vector length	Mean error (re: controls)	Sample size
(1) Control:			
Wire loft	0.39	0°	210
(2) Glass loft	0.14	34°	774
(3) Louvred loft	0.24	28°	569
(4) Glass and open loft	0.24	141°	289
(5) Wood and louvered loft	0.22	177°	521

Data calculated from ref. 48.

is based on the gradual and basically regular change in the dip angle and strength of the Earth's magnetic field. The field strength of the Earth at 40°N latitude in the US is about 0.5 G (50,000 nT). The field has several sources: by far the strongest is generated by the Earth's core, while a much weaker effect arises inductively from the movement of ions around the Earth by the jet streams. Because the daily heating and cooling of the atmosphere displaces the jet streams north and south, a more-or-less regular circadian variation of 30–50 nT is observed on the ground (Fig. 4). The total magnetic field of the Earth in the northeastern United States increases at about 3–5 nT km⁻¹ to the NNW, the vertical strength rises roughly twice as fast, and the dip angle increases at about 0.01° per km.

To account for the accuracy displayed by pigeons wearing frosted contact lenses⁷, a seemingly impossible sensitivity of 10–30 nT or 0.03° would be necessary; but just such sensitivities can be inferred from the behaviour of birds. After solar flares, for example, enormous numbers of extra ions appear in the jet streams, causing irregular changes in the strength of the Earth's field ranging from 10 to 1,000 nT (Fig. 4). These 'storms' have roughly dosage-dependent effects on birds⁸⁴. For example, the release-site bias of Ithaca pigeons released at Campbell, New York, is rotated up to 40° counter-clockwise (in addition to the usual 12° bias) during moderate storms (Fig. 5a). (Regrettably, no data on how scattered the birds' vanishing bearings were during such storms has been published.) A retrospective analysis of 18 years of homing pigeons races in Italy^{85,86} confirms a strong negative correlation between homing speed and the sunspot activity which is responsible for such storms, and Yeagley pointed out a similar relationship in American races⁴². Similarly, Moore observed that the increasing levels of magnetic storm activity increased the angular deviation of free-flying nocturnal migrants⁸⁷, while Southern found a substantial effect on the orientation of gull chicks⁸⁸. These various observations imply a sensitivity of 10–30 nT. Since even a 1,000 nT storm could not, in the most extreme conditions, rotate a compass needle even 2°, and since the effects on pigeons and gulls are observed when the sun is clearly visible, the phenomenon is more likely to be related to the map sense.

Although the effects of solar activity on pigeon orientation are well documented and intriguing, naturally occurring variations in magnetic topography provide a static and therefore more convenient index to magnetic sensitivity. To use a magnetic map, pigeons would have to extrapolate the local rate and direction of magnetic gradients (inferred either during exercise flights or during transport to the release site) in order to judge the distance and direction of displacement. Hence any slight nonlinearity in values at the loft ought to result in a vague but

systematic pattern of errors at release sites (imposed, of course, on any site-specific deviations) roughly aligned with the general direction of the regional magnetic gradients. This seems to be the case. Figure 1b, for example, reveals a pronounced tendency for site biases around Ithaca to be clockwise east of the NNW line of magnetic field gradients, and counter-clockwise to the west. It is as though 'pigeon' Ithaca is NNW of geographic Ithaca. Similarly, the pattern of biases around Schmidt-Koenig's former loft at Frankfurt (where the gradient runs west) and several locations mapped by Wallraff consistently fall into this pattern. As would be expected in some cases (one of Wallraff's as well as Walcott's loft in Lincoln, Massachusetts), the clockwise/counter-clockwise pattern is reversed as though the loft is magnetically south of its true location along the axis of the regional field gradients.

The Earth's magnetic field also has permanent localized anomalies, generally created by areas of high magnetic permeability such as iron deposits. These provide a favourable path for magnetic lines of force, concentrating the lines and thereby raising the local field strength. Walcott has carried out an extensive series of releases at anomalies of various strengths and finds a clear relationship between the magnitude of the anomaly and the degree of scatter in vanishing bearings (Fig. 5b)⁸⁹. Tracking of birds from normal (Fig. 6a) and anomalous sites (Fig. 6b) on sunny days illustrates the dramatic and long-lasting nature of the effect. It seems clear that these small irregularities in magnetic field strength leave birds with almost no idea of where they are; and the persistence of the effect for a time after they have escaped from the anomaly suggests a time-averaging processing strategy similar to that of bees⁹¹. It is also worth noting that the 10–30 nT level at which effects begin to be evident corresponds to the predicted threshold for a magnetic map sense based on the accuracy displayed by the birds wearing frosted lenses⁷.

Along the same lines Wagner^{92,93} has found a strong tendency for pigeons to respond to magnetic slopes of about 30 nT km⁻¹. His birds flew down slopes regardless of home direction—a sensible strategy since gradients of more than 5 nT km⁻¹ indicate an anomaly—and, since virtually all anomalies are the result of elevated local field strength, down is almost always the way out to the magnetic 'plains' where more reliable map readings would be possible (see Fig. 6). (Similar though less well documented effects of magnetic topography have been reported by both Talkington⁹⁴ and Graue⁹⁵.) Moreover, such a 'fly-down' strategy could explain the mysterious difficulty of nearly all Ithaca pigeons to orient accurately at the Jersey Hill fire tower in western New York state. The few birds which have been tracked from aircraft wander about within a few kilometres tracing out the limits of what is, to a first approximation, a magnetic 'valley'.

Further suggestive evidence in favour of a magnetic map sense comes from three diverse sets of experiments by Kiepenheuer⁹⁶, Wilschko's group⁹⁷, and even Papi and his colleagues⁹⁸. In each case an alteration of the magnetic field felt by the birds on the way to the release site resulted in a significant change in orientation. These recent results are in marked contrast to older experiments (cited earlier) with less completely controlled fields which found no effect. The exact reason for these differences is not clear.

Problems with the magnetic hypothesis

Encouraging as all these data seem, the magnetic map hypothesis faces severe problems. The first is that two axes are needed for bicoordinate navigation. The magnetic parameters available are total field strength, vertical strength, horizontal strength, declination (the deviation of the horizontal vector from true north), and inclination (the dip angle). Declination is useless since it requires knowing true north at the release site even under overcast skies; and both inclination and total field strength use up the same two components: vertical and horizontal strength. Hence, birds would have either to compare horizontal against vertical or total field strength which would form a grid in the northeastern United States intersecting at

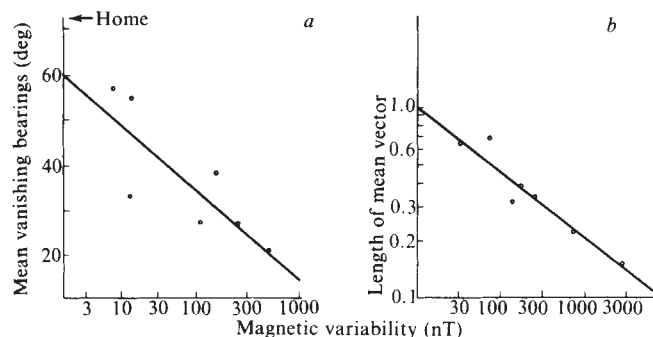


Fig. 5 Effect of small magnetic field disturbances on homing pigeons. *a*, Birds released at Campbell, New York normally show a site bias of 12° CCW. This bias is exaggerated during magnetic storms. Storm data represents a 12-hour average of the most active magnetic component as recorded at Fredricksburg. Data from Keeton *et al.*⁸⁴. *b*, Vanishing bearings of birds released at magnetic anomalies are more scattered in proportion to the strength of the anomaly. The data, from Walcott⁸⁹, are based on the magnetic variability in the homeward direction over the first kilometre.

about 20°. If pigeons were using such a grid, then departures ought to be more accurate and less scattered to the ENE and WSW. In fact, only a slight trend of this sort is apparent in the data. Of the two other possibilities, total intensity and dip angle (the ratio of vertical and horizontal intensities), only the latter seems likely: there are areas with large, gentle variations in field intensities—a result probably of either weak, large-scale differences in ground permeability to magnetic fields, or strong but deep anomalies—which would overwhelm any simplistic

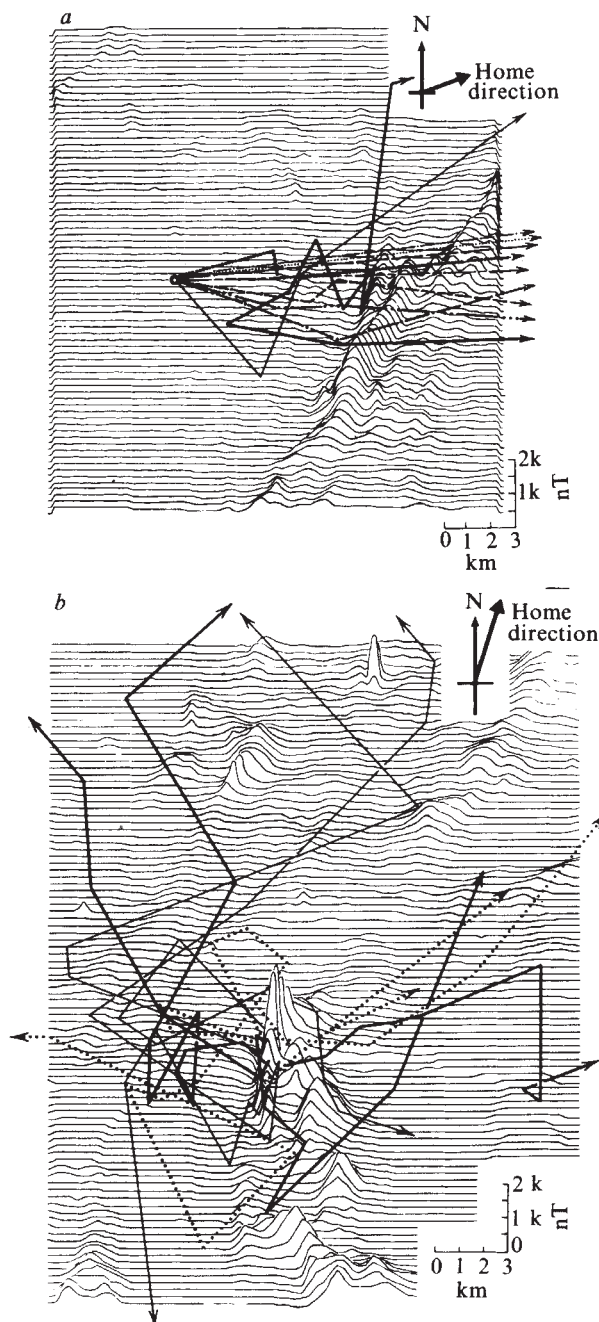


Fig. 6 Flight paths of pigeons away from release sites. Total magnetic intensity (the only parameter available from the aeromagnetic surveys by the US Geological Survey upon which these illustrations are based) increases steadily at 4 nT km^{-1} to the NNW, creating an inclined plane. Only the anomalies in this pattern have been drawn here and are seen as 'hills'. *a*, Pigeons released from a magnetically normal site near Worcester, Massachusetts depart directly away, displaying the slight clockwise bias typical of this site. *b*, Birds released at the magnetic anomaly at Iron Mine Hill, Rhode Island, on the other hand, are thoroughly disoriented. At this site the field strength is approximately 6% high. Data from (a) Walcott (personal communication) and (b) Walcott⁸⁹; drawings from Gould⁹⁰.

attempt to make sense of regional patterns of total field intensity but which might not affect the horizontal/vertical ratio. This ratio, the inclination or dip, is precisely the measure employed for magnetic compass orientation in robins³¹ and probably pigeons³⁰. But if birds must abandon total intensity for dip, they must be able to measure true vertical—presumably gravity—as a reference angle with great precision (that is, to within 0.03° , a difficult but not impossible task). Perhaps this could help explain some of the puzzling, weak gravitational effects reported by the Keeton group⁹⁹. To further confound matters, most magnetic anomalies, created as they are by iron deposits, are gravity anomalies as well.

Another apparent problem is that strong magnets and Helmholtz coils normally do not destroy orientation on sunny days^{29,30,100}. How could pigeons possibly measure such minute field changes in the face of strong static fields? Interestingly, however, the vanishing bearings of pigeons given such strong-field treatments correspond to those seen on days of severe magnetic storms¹⁰¹—suggesting a map effect. Moreover, when Walcott released magnet-equipped pigeons at anomalies, the effect of those small geographical disturbances persisted¹⁰², indicating that these small perturbations were being perceived through the strongly altered background. Taken together, both the data and theoretical considerations suggest that any magnetic map detector must attend solely to the small changes between and around both loft and release site, and ignore the far larger but utterly useless background field which is essential for the cloudy day compass system. A further constraint is that such a map system must have a fairly long time constant since (1) pigeons can usually fly through anomalies with little effect, (2) the birds are not disoriented by wing-mounted magnets which would create an alternating background field^{38,39,103}, (3) they seem less disturbed by the brief but powerful 'pulses' in magnetic field storms than by longer-term changes in background field, and (4) homing pigeons do not recover from releases at anomalies until they have been out of the field for some time (Fig. 6*b*).

A final problem for any magnetic map system is how the nervous system could possibly measure such small intensity variations. Although Kirschvink and Gould have shown that the vast numbers of single domains of magnetite discovered in pigeons could produce a resolution better than 1 nT ¹⁰⁴, whether either these calculations or the magnetite have anything to do with pigeon navigation remains to be seen.

The real difficulties with the magnetic map hypothesis, then, are threefold. (1) It has no satisfactory explanation of the necessary second coordinate (it could be magnetism again, odour, infrasound, the Sun (though obviously not on cloudy days), or even Coriolis force). (2) It is based almost entirely on correlations rather than direct, controlled experimental manipulations (the effects of strong, artificial field alterations during transport, for example, ought to have generally unpredictable or even negligible effects). And (3) there is as yet no convincing explanation of either the sensory basis of magnetic field perception or the processing of that data to extract map information. Given the evidence for both olfactory and magnetic effects, it would be attractive to suppose that *both* answers are correct: perhaps, by analogy with the solar and magnetic compass of pigeons, birds can use either as the situation demands; or, both might be used simultaneously. If the first possibility were correct, we would expect that repeated releases at magnetic anomalies ought to force pigeons into using the alternative olfactory map system. However, since Walcott has released the same pigeons repeatedly at different anomalies, and such birds simply do not improve, the inference is that birds must lack an alternative system. Of the possibility that olfactory and magnetic information must be used together, one can only say that it seems at least as plausible as that they use magnetic information alone or with some unknown factor.

It is ironic that, more than three decades after Yeagley's original hypothesis, we are still largely in the dark and, for all we know, his explanation may yet turn out to have been rather close

to the mark. We probably now have nearly all the pieces of the puzzle before us (and, doubtless, several utterly irrelevant ones as well) but for the immediate future, at least, the nature of the animal map sense seems likely to retain its status as the most

elusive and intriguing mystery in animal behaviour.

I thank C. Walcott, C. G. Gould and D. R. Griffin for their contributions, and the NSF for its support through grant BNS 78-24754.

1. Elsner, B. in *Animal Orientation, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 194–198 (Springer, Berlin, 1978).
2. Windsor, D. M. *Anim. Behav.* **23**, 334–343 (1975).
3. Keeton, W. T. *J. comp. Physiol.* **86**, 1–16 (1973).
4. Wallraff, H. G. *Z. Tierpsychol.* **27**, 303–351 (1970).
5. Wallraff, H. G. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 171–183 (Springer, Berlin, 1978).
6. Michener, M. & Walcott, C. *J. exp. Biol.* **47**, 99–131 (1967).
7. Schmidt-Koenig, K. & Walcott, C. *Anim. Behav.* **26**, 480–486 (1978).
8. Keeton, W. T. *Sci. Am.* **231** (6), 96–107 (1974).
9. Gould, J. L. *Amer. Sci.* **68**, 256–267 (1980).
10. Griffin, D. R. *Bird-Banding* **14**, 7–33 (1943).
11. Wallraff, H. G. *J. comp. Physiol.* **139**, 193–201 (1980).
12. Wallraff, H. G. *J. comp. Physiol.* **139**, 203–208 (1980).
13. Kluijver, H. W. *Ardea* **24**, 227–239 (1935).
14. Walcott, C. & Schmidt-Koenig, K. *Auk* **90**, 281–286 (1973).
15. Hachet-Souplet, P. *Rev. Sci.* **29**, 231–238 (1911).
16. Huizinger, E. *Pflügers Arch. ges. Physiol.* **236**, 52–58 (1935).
17. Wallraff, H. G. *Z. vergl. Physiol.* **50**, 313–330 (1965).
18. Wallraff, H. G. *Nature* **236**, 223–224 (1972).
19. Keeton, W. T. *Adv. Study Behav.* **5**, 47–132 (1974).
20. Kramer, G. *Verh. Dcut. Zool. ges. Freiburg* **1952**, 72–84 (1953).
21. Kramer, G. & St. Paul, U. v. *Naturwissenschaften* **37**, 526–527 (1950).
22. Kramer, G. & Riese, E. *Z. Tierpsychol.* **9**, 245–251 (1952).
23. Hoffmann, K. *Z. Tierpsychol.* **11**, 453–475 (1974).
24. Schmidt-Koenig, K. *Z. Tierpsychol.* **18**, 221–244 (1961).
25. Keeton, W. T. *Science* **165**, 922–928 (1969).
26. Schmidt-Koenig, K. in *Animal Orientation and Navigation* (eds Galler, S. R. et al.) 275–282 (US Govt Printing Office, Washington, D.C., 1972).
27. Merkel, F. W. & Fromme, H. G. *Naturwissenschaften* **45**, 499–500 (1958).
28. Merkel, F. W. & Wiltshko, W. *Vogelwarte* **23**, 71–77 (1965).
29. Keeton, W. T. *Proc. natn. Acad. Sci. U.S.A.* **68**, 102–106 (1971).
30. Walcott, C. & Green, R. P. *Science* **184**, 180–182 (1974).
31. Wiltshko, W. & Wiltshko, R. *Science* **176**, 62–64 (1972).
32. Visalberghi, E. & Alleva, E. *Biol. Bull.* **156**, 246–256 (1979).
33. Wiltshko, W., Nohr, D. & Wiltshko, R. *Science* **214**, 343–345 (1981).
34. Wiltshko, W., Wiltshko, R. & Keeton, W. T. *Behav. Ecol. Sociobiol.* **1**, 229–243 (1976).
35. Walcott, C., Gould, J. L. & Kirschvink, J. L. *Science* **205**, 1027–1029 (1979).
36. Schreiber, B. *Boll. Zool.* **48**, 101 (1981).
37. Wiltshko, W. & Wiltshko, R. *Z. Tierpsychol.* **37**, 337–355 (1975).
38. Wiltshko, W. & Wiltshko, R. *J. comp. Physiol.* **109**, 91–99 (1976).
39. Bingman, V. P. (in preparation).
40. Yeagley, H. L. *J. appl. Phys.* **18**, 1035–1063 (1947).
41. Yeagley, H. L. *J. appl. Phys.* **22**, 746–760 (1951).
42. Schmidt-Koenig, K. *Avian Orientation and Navigation*, 131–132 (Academic, New York, 1979).
43. Matthews, G. T. V. *J. exp. Biol.* **30**, 243–267 (1953).
44. Matthews, G. T. V. *Bird Navigation* (Cambridge University Press, 1955).
45. Kramer, G. *Ibis* **101**, 399–416 (1959).
46. Wallraff, H. G. *Z. vergl. Physiol.* **52**, 215–259 (1966).
47. Wallraff, H. G. *Z. vergl. Physiol.* **68**, 182–201 (1970).
48. Wallraff, H. G. *Behav. Ecol. Sociobiol.* **5**, 201–225 (1979).
49. Papi, F. *Verh. Dcut. Zool. Ges.* **1976**, 184–205 (1976).
50. Papi, F., Fiore, L., Fiaschi, V. & Benvenuti, S. *Monit. Zool. Ital.* **6**, 85–95 (1972).
51. Benvenuti, S., Fiaschi, V., Fiore, L. & Papi, F. *J. comp. Physiol.* **83**, 81–92 (1973).
52. Papi, F., Mariotti, G., Foà, A. & Fiaschi, V. *J. comp. Physiol.* **135**, 227–232 (1980).
53. Papi, F., Fiaschi, V. & Benvenuti, S. *Monit. Zool. Ital.* **7**, 129–133 (1973).
54. Papi, F., Ioalé, P., Fiaschi, V., Benvenuti, S. & Baldaccini, N. E. *J. comp. Physiol.* **94**, 187–193 (1974).
55. Baldaccini, N. E., Benvenuti, S., Fiaschi, V. & Papi, F. *J. comp. Physiol.* **99**, 177–186 (1975).
56. Baldaccini, N. E., Benvenuti, S., Fiaschi, V., Ioalé, P. & Papi, F. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 78–91 (Springer, Berlin, 1978).
57. Benvenuti, S., Fiaschi, V. & Foà, A. *J. comp. Physiol.* **120**, 173–179 (1977).
58. Ioalé, P., Papi, F., Fiaschi, V. & Baldaccini, N. E. *J. comp. Physiol.* **128**, 285–295 (1978).
59. Schmidt-Koenig, K. & Phillips, J. B. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 119–124 (Springer, Berlin, 1978).
60. Kreithen, M. L. & Eisner, T. *Nature* **272**, 347–348 (1978).
61. Yodlowski, M. L., Kreithen, M. L. & Keeton, W. T. *Nature* **265**, 725–726 (1977).
62. Kreithen, M. L. & Keeton, W. T. *J. comp. Physiol.* **89**, 73–82 (1974).
63. Kreithen, M. L. & Keeton, W. T. *J. comp. Physiol.* **89**, 83–92 (1974).
64. Papi, F., Keeton, W. T., Brown, A. I. & Benvenuti, S. *J. comp. Physiol.* **128**, 303–317 (1978).
65. Benvenuti, S. *Z. Tierpsychol.* **51**, 406–414 (1979).
66. Wallraff, H. G. *J. comp. Physiol.* **139**, 209–224 (1980).
67. Keeton, W. T., Kreithen, M. L. & Hermayer, K. L. *J. comp. Physiol.* **114**, 289–299 (1977).
68. Kiepenheuer, J. *Behav. Ecol. Sociobiol.* **6**, 11–22 (1979).
69. Keeton, W. T. *Monit. Zool. Ital.* **8**, 227–234 (1974).
70. Hartwick, R., Kiepenheuer, J. & Schmidt-Koenig, K. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 107–118 (Springer, Berlin, 1968).
71. Keeton, W. T. & Brown, A. I. *J. comp. Physiol.* **105**, 259–266 (1976).
72. Kiepenheuer, J. *Behav. Ecol. Sociobiol.* **3**, 393–395 (1978).
73. Waldvogel, J. A., Benvenuti, S., Keeton, W. T. & Papi, F. *J. comp. Physiol.* **128**, 297–301 (1978).
74. Waldvogel, J. A., Phillips, J. B., McCorkle, D. R. & Keeton, W. T. *Behav. Ecol. Sociobiol.* **7**, 207–211 (1980).
75. Waldvogel, J. A. thesis, Cornell Univ. (1981).
76. Brines, M. L. *J. theor. Biol.* **86**, 507–512 (1980).
77. Brines, M. L. thesis, Rockefeller Univ. (1978).
78. Brines, M. L. & Gould, J. L. *Science* **206**, 571–573 (1979).
79. Brines, M. L. & Gould, J. L. *J. exp. Biol.* (in the press).
80. Able, K. P. (in preparation).
81. Baldaccini, N. E., Benvenuti, S., Fiaschi, V., Ioalé, P. & Papi, F. *J. comp. Physiol.* **94**, 85–96 (1974).
82. Walcott, C. *IEEE Trans. Magnetics* **16**, 1008–1013 (1980).
83. Moore, B. *Nature* **285**, 69–70 (1980).
84. Keeton, W. T., Larkin, T. S. & Windsor, D. M. *J. comp. Physiol.* **95**, 95–103 (1974).
85. Schreiber, B. & Rossi, O. *Boll. Zool.* **43**, 317–320 (1976).
86. Schreiber, B. & Rossi, O. *IEEE Trans. Magnetics* **14**, 961–963 (1978).
87. Moore, F. R. *Science* **196**, 682–684 (1977).
88. Southern, W. E. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 311–317 (Springer, Berlin, 1978).
89. Walcott, C. in *Animal Migration, Navigation and Homing* (eds Schmidt-Koenig, K. & Keeton, W. T.) 143–151 (Springer, Berlin, 1978).
90. Gould, J. L. *Ethology: the Mechanisms and Evolution of Behavior* (Norton, New York, 1982).
91. Gould, J. L. *Science* **207**, 545–547 (1980).
92. Wagner, G. *Rev. Suisse Zool.* **83**, 883–890 (1976).
93. Frei, U. & Wagner, G. *Rev. Suisse Zool.* **83**, 891–897 (1976).
94. Talkington, L. *Am. Zool.* **7**, 199 (1967).
95. Graue, L. C. *Am. Zool.* **5**, 704 (1965).
96. Kiepenheuer, J. in *Animal Migration, Navigation and Orientation* (eds Schmidt-Koenig, K. & Keeton, W. T.) 135–142 (Springer, Berlin, 1978).
97. Wiltshko, R., Wiltshko, W. & Keeton, W. T. in *Animal Migration, Navigation and Orientation* (eds Schmidt-Koenig, K. & Keeton, W. T.) 152–161 (Springer, Berlin, 1978).
98. Papi, F., Ioalé, P., Fiaschi, V., Benvenuti, S. & Baldaccini, N. E. in *Animal Migration, Navigation and Homing*, 65–77 (Springer, Berlin, 1978).
99. Larkin, T. S. & Keeton, W. T. in *Animal Migration, Navigation and Orientation*, (eds Schmidt-Koenig, K. & Keeton, W. T.) 92–106 (Springer, Berlin, 1978).
100. Walcott, C. *J. exp. Biol.* **70**, 105–123 (1977).
101. Larkin, T. S. & Keeton, W. T. *J. comp. Physiol.* **110**, 227–231 (1976).
102. Walcott, C. *J. exp. Biol.* **86**, 349–352 (1980).
103. Matthews, G. T. V. *J. exp. Biol.* **30**, 243–267 (1955).
104. Kirschvink, J. L. & Gould, J. L. *Biosystems* **13**, 181–201 (1981).

ARTICLES

Pancake detonation of stars by black holes in galactic nuclei

B. Carter & J. P. Luminet

Groupe d'Astrophysique Relativiste, Observatoire de Paris, 92190 Meudon, France

Recent efforts to understand exotic phenomena in galactic nuclei commonly postulate the presence of a massive black hole accreting gas produced by tidal or collisional disruption of stars. For black holes in the mass range 10^4 – $10^7 M_\odot$, individual stars penetrating well inside the Roche radius will undergo compression to a short-lived pancake configuration very similar to that produced by a high velocity symmetric collision of the kind likely to occur in the neighbourhood of black holes in the higher mass range $\geq 10^9 M_\odot$. Thermonuclear energy release ensuing in the more extreme events may be sufficient to modify substantially the working of the entire accretion process.

MANY of the most plausible models for explaining spectacular energy release phenomena in galactic nuclei are based on the release of gravitational energy by gas accreting on to a very massive central black hole¹. A likely mechanism² for providing the gas necessary to fuel such a model would be the breakup of ordinary stars, either by the Roche tidal effect in the case of a black hole with mass M below the Hills³ limit, $M \leq 10^8 M_\odot$, or else by the effect of collisions in the case of a more massive black hole with a sufficiently dense surrounding star cluster. Gas release by tidal or collisional breakup (as well as by mechanisms such as supernovae and ordinary stellar winds) may also be important at other relatively quiescent stages in the history of a galactic nucleus containing a black hole, that is at times when the most observationally spectacular phenomena (such as quasars) are absent.

This article draws attention to a neglected effect which could be significant in many situations where tidal or collisional disruption occurs, namely that in the more extreme events (which may constitute a not negligible fraction of the total) the disruption will be preceded by a short-lived phase of high compression to a roughly pancake shaped configuration in which the density and temperature may rise enough to detonate effectively some significant fraction of the available thermonuclear fuel.

A very flat pancake type configuration will be formed briefly by any nearly head-on collision between roughly similar stars at very high relative velocities (analogous pancake formation being experimentally familiar in the microscopic context of very high energy proton-proton collision). We have in mind in particular the kind of event that is likely to occur near a black hole with the very large mass $M \geq 10^9 M_\odot$ that is thought⁴ to be necessary to account for the most extreme quasar phenomena. Beyond 0.1 pc from such an object, typical stellar collisions will have velocities $\geq 10^4 \text{ km s}^{-1}$, thus exceeding by a factor $\beta \geq 10$ the minimum of the order of 10^3 km s^{-1} needed for disruption. Detailed studies of high velocity ($\geq 10^4 \text{ km s}^{-1}$) collisions will require an elaborate hydrodynamic treatment, with allowance for shocks, of the kind already available⁵ for the intermediate velocity ($\approx 10^3 \text{ km s}^{-1}$) collisions in which a moderate degree of flattening is already present. However, a preliminary idea of the effects to be expected may be obtained from consideration of the simpler but otherwise very similar configuration that arises during tidal disruption of a single star, whose evolution can be followed approximately^{6,7} in terms only of ordinary differential equations using an adiabatic affine star model which, although highly idealized, can, nevertheless, be expected to provide a qualitatively valid description of the behaviour of the stellar core before, if not after the instant of maximum compression.

The phenomena of tidal disruption of a self gravitating body passing within the Roche radius

$$R_R \approx M^{1/3} \rho_*^{-1/3} \quad (1)$$

(where ρ_* is its characteristic central density) has been studied for over a century, but mainly in terms of an incompressible fluid model. It does not seem to have been realized that adequate allowance for compressibility leads to the prediction that a star penetrating deeply within the Roche radius will pass through a phase of compression to a highly flattened pancake configuration, closely similar, in both spatial and temporal characteristics, to that produced by a symmetric collision involving a pair of stars. It turns out that the effect of a collision with stellar velocities exceeding the central characteristic velocity (which is a few hundreds of km s^{-1} for a typical main sequence star) by a factor of the order of β can be roughly simulated by a single star following an orbit that penetrates to within a pericentre radius R_p given in order of magnitude by

$$\beta \approx R_R / R_p \quad (2)$$

(a value β of the order of 10 thus being sufficient to reproduce the effect of a collision with relative velocity of the order of 10^4 km s^{-1}). Not only can the study of thermonuclear reactions in such a tidally produced pancake give a first rough indication

of what may occur in a high velocity stellar collision, but it would also appear to be an astrophysically interesting phenomenon in its own right in the context of a black hole in the moderate mass range $10^4 M_\odot$ to $10^7 M_\odot$ for which it is possible to obtain a large value of β without the star either being penetrated by or swallowed by the black hole. The potential importance of this phenomenon stems from the fact that deep penetration is much less improbable than would be expected from purely geometric considerations. Just as the probability of high velocity collisions deep in the potential well of an $M \geq 10^9 M_\odot$ black hole is enhanced far above what would be proportional to the corresponding volume (by an amount whose exact calculation is difficult) in consequence of the well known cusp effect in the stellar density distribution (see ref. 8), so, similarly, the probability of penetration of an individual star deep within the tidal field of a more moderate sized black hole will also be much higher than would be deduced from the corresponding purely geometrical cross-section, by an amount that can be relatively easily estimated from the characteristics of the individual approximately parabolic orbit. We conclude that the fraction of all tidally disrupted stars with penetration factor exceeding a given value β will be of the order of magnitude of β^{-1} .

Extent and duration of compression

To derive equation (2) it is sufficient to see that the tidal force on the star (which is inversely proportional to the cube of the radial distance R from the hole) will rapidly dominate the internal pressure and self-gravitational forces on the star as it begins to penetrate substantially within the Roche radius R_R (the latter being characterized by the condition that the tidal and internal forces have comparable magnitude). In these circumstances it will become a good approximation to treat the individual particles of the star as undergoing free fall in the external gravitational field on orbits of approximately parabolic form in planes all passing very nearly through the perihelion point on the centre of mass trajectory. These planes will be confined roughly within an angle

$$\alpha \approx R_*(R_p R_R)^{-1/2} \quad (3)$$

where R_* is the characteristic radius of the star in its approximately unperturbed state before passage across the Roche radius, as given in terms of the stellar mass M_* by

$$R_* \approx M_*^{1/3} \rho_*^{-1/3} \quad (4)$$

Thus in the centre of mass frame of the star the particles will effectively be accelerated towards the orbital plane, with maximum velocity, u say, of the order of αv where v is the orbital velocity at the pericentre, that is

$$u \approx \alpha \left(\frac{GM}{R_p} \right)^{1/2} \quad (5)$$

Hence we immediately see that the relative velocity u of what may be described as the collision of the star with itself will have the required form

$$u \approx \beta \left(\frac{GM_*}{R_*} \right)^{1/2} \quad (6)$$

where the factor $(GM_*/R_*)^{1/2}$ is interpretable as the characteristic velocity of sound in the core of the star.

When matter becomes sufficiently highly compressed, the free-fall treatment will, of course, cease to be valid: as the mass centre passes through the pericentre, the tidal forces, although high, will be overtaken by a very sharp last minute rise in the pressure forces, which will cause the star to bounce back towards a more isotropic configuration. Note that stellar deformation in directions in the plane of the orbit will not have had time to become important at the instant of passage through the pericentre (the 'tube of toothpaste' extrusion effect that is familiar from the incompressible case cannot have developed significantly at this stage) so the maximum degree of compression of the resulting pancake can be estimated directly from the

requirement that the kinetic energy of the 'self-collision' be entirely converted into internal energy of the gas at the instant of turn around. If this energy is predominantly thermal, as will be the case in main sequence stars, the maximum central temperature Θ_m in the pancake will be given in terms of its value Θ_* in the unperturbed state of the star by

$$\Theta_m \approx \beta^2 \Theta_* \quad (7)$$

This formula will also be valid when the predominant internal energy is that of a non-relativistic degenerate electron gas, which scales proportionally to the non-degenerate thermal energy of the gas of positive ions. As long as the gas remains non-relativistic the corresponding maximum central density ρ_m will be given by

$$\rho_m \approx \beta^3 \rho_* \quad (8)$$

and the duration τ_m of the phase of maximum compression will be given by

$$\tau_m \approx \beta^{-4} (G\rho_*)^{-1/2} \quad (9)$$

where $(G\rho_*)^{-1/2}$ is the free-fall time scale of the star (which is automatically the same as the orbital time scale at the Roche radius R_R) whose value is of the order of 10^3 s for small main sequence stars.

Although the above arguments are crude we have confirmed them (see Fig. 1) by explicit numerical integration of the equations of motion governing an affine star model (in which the constant density layers are treated as retaining an ellipsoidal form) which represents the simplest and most natural extension of the incompressible fluid model (on which most previous numerical studies of tidal forces have been based). While probably at least qualitatively adequate during the compressive phase, such a treatment can be expected to be invalidated during the phase of subsequent expansion by the effect of shocks, whose potential importance in the envelope (as opposed to the central regions with which we are concerned here) has been pointed out by Lidskii and Ozernoy⁹, who seem to have been the first to appreciate the significance of the tidal field's initial tendency to cause compression.

Thermonuclear detonation

It is evident from equation (7) that a quite modest Roche penetration factor $\beta \approx 10$ can raise the temperature of a main sequence star from its usual value $\Theta_* \approx 10^7$ K to near the value $\Theta_m \approx 2 \times 10^9$ K at which (according to the standard formula given for example by Fowler *et al.*¹⁰) helium combustion by the triple- α reaction proceeds at a maximum rate for a given density, the relevant time scale being given (in c.g.s.) by

$$\tau_\alpha \approx 10^{11} (\rho_m X_\alpha)^{-2} \quad (10)$$

where the helium mass fraction X_α will have at least the cosmological value $X_\alpha \approx 1/4$. As the energy that can be obtained from helium combustion is only of the same order as the thermal energy (borrowed temporarily from the gravitational field) already present, the condition $\tau_\alpha \leq \tau_m$ is all that is required for a substantial fraction of this energy to be released. As the values $X_\alpha \approx 1/4$ and $\beta \approx 10$ lead roughly to $\tau_m/\tau_\alpha \approx 10^{-5} \rho_*^{3/2}$, we see that the efficiency of helium detonation depends primarily on the central density of the star, being reasonably high only when ρ_* approaches the order of 10^3 g cm⁻³. By coincidence this is roughly the upper limit for main sequence central densities, being obtained near their lower mass limit, $M_* \approx 10^{-1} M_\odot$.

The foregoing leads us to conclude that the most favourable conditions for efficient helium detonation will occur in conceivably very numerous (see ref. 11) 'dark dwarf' stars that are just too small to reach the main sequence but which, nevertheless, maintain central temperatures $\geq 10^6$ at densities ranging up to 10^3 g cm⁻³ during time scales of the order of 10^9 yr (see ref. 12). To reach $\Theta_m \approx 10^9$ K such a star needs perhaps twice as large a penetration factor β as in the main sequence case (with a correspondingly halved probability), but as a result

ρ_m will be raised from a main sequence maximum of the order of 10^6 g cm⁻³ to a value that can exceed 10^7 g cm⁻³. At such densities the degenerate electrons will enter the relativistic regime, so that equations (7)–(9) will need modification, but allowance for this seems to confirm the conclusion that τ_m/τ_α attains a value of the order of 1.

The carbon-12 formed in such a detonation process will immediately undergo further processing by α -capture and proton-capture reactions (whose details depend very sensitively on M_* and β) so that the total energy release may exceed the internal binding energy of the star by a factor of the order of 10^2 . Thus instead of forming a diffuse cloud weakly bound to the black hole as in the non-violent disruption scenario described by Hills³, the gas will be ultimately liberated from the star in the form of a cloud expanding at a rate far in excess of the velocity of escape from the black hole. Any discussion of detailed consequences is beyond the scope of this article. We mention only that the immediate effect of the pancake detonation will be to accelerate the matter orthogonally to the orbital plane, but that subsequent radioactive energy release may lead to slingshot ejection of a fraction of the matter at very high velocities in the direction opposite to that along which the star originally approaches the hole, and to the corresponding capture by the black hole of comparable fraction of the matter in very tightly bound cloud or disk, with thermal or orbital speeds $> 10^4$ km s⁻¹. It is even conceivable that a self-sustaining hot C–N–O or rapid proton capture¹³ hydrogen burning process may be set off in the expanding cloud, with even more spectacular consequences, but only a careful hydrodynamic treatment with adequate treatment of shocks will be able to check this. Analysis of the overall scenario resulting from interaction of gas from many disrupted stars will be even more complicated but we can conclude that the likely detonation of a fraction of the order of 1% (arising from say a 10% fraction with sufficiently high penetration factor β , of which a less easily estimable fraction, perhaps also of the order of 10%, may lie in the susceptible mass range) will probably be sufficient to modify significantly the overall energetics of the entire gas release process. (Analogous conclusions are probably valid for the more complicated situation resulting from collisional, as opposed to tidal, disruptions in the neighbourhood of much more massive black holes: the unfavourable statistical factor resulting from the probable necessity that both participants in the collision should lie in the susceptible mass range seems likely to be compensated by the enhanced probability of very high velocities due to the cusp effect.)

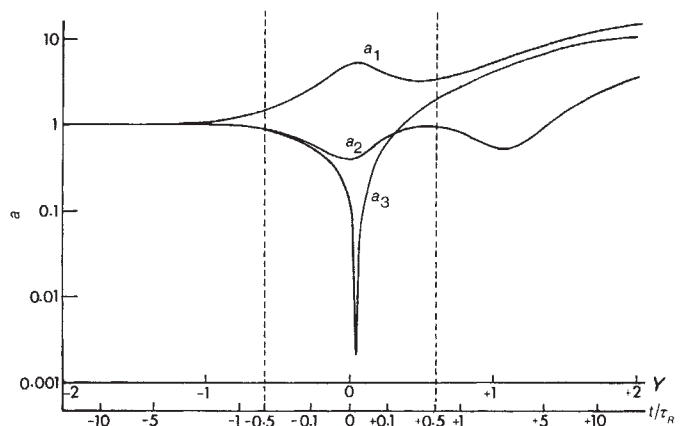


Fig. 1 The numerically calculated magnitudes of the principal axes a_1 , a_2 , a_3 (a_3 being orthogonal to the orbital plane) of a polytropic ($\gamma = 5/3$) affine (that is, ellipsoidally deformed) star model are plotted on a logarithmic scale as a function of the distance Y from the axis of a parabolic trajectory with Roche penetration factor $\beta = 10$. The corresponding time t , normalized as a fraction of the stellar oscillation time scale $\tau_R = (G\rho_*)^{-1/2}$, is indicated on a separate scale below, the approximate moments at which the star enters and leaves the Roche radius R_R being indicated by dotted lines.

Although the triple- α reaction we have been considering provides the only opening for significant thermonuclear release from the compression of population II stars (consisting almost exclusively of hydrogen and helium) it has been drawn to our attention that if the stellar population in the galactic nucleus has composition similar to that of population I stars, such as the Sun, then energy release on an individually more modest scale, but in a much larger fraction of the disrupted stars, can result from C–N–O cycle reactions. It is not possible to carry through a complete C–N–O cycle during the short lifetime of the pancake configuration because some of the steps inevitably involve slow weak decays. Nevertheless, the first highly temperature sensitive proton-capture stages (together with the later disintegration of the products) can liberate several times the binding energy of a star of solar type, if the temperature in the pancake rises above a few times 10^8 K. As this effect is much less density sensitive and requires a lower Roche penetration

factor ($\beta \approx 5$ would suffice) than the triple- α reaction, it will occur in a much larger fraction of the stars involved. Hence, although the individual events are much less spectacular, the total energy release from this process can (when averaged over all disrupted stars) provide a contribution to the overall energetics that is no less important than that arising from helium detonation. Indeed for a black hole near the Hills mass limit (where the larger penetration factors required for triple- α detonation are not possible) energy release from proton-capture will be the dominant contribution. The net effect will presumably be to convert the gas from the disrupted stars into an outgoing wind, leaving only a small fraction to be accreted by the hole.

We thank many colleagues, particularly S. Bonazzola, S. Collin, T. Damour, D. Péquignot, M. Rees, J. Schneider, H. Sol and C. Zaidins, for helpful discussions.

Received 17 September 1981; accepted 27 January 1982.

1. Rees, M. *Phys. Scr.* **17**, 193 (1978).
2. Frank, J. & Rees, M. *Mon. Not. R. astr. Soc.* **176**, 683 (1976).
3. Hills, J. G. *Nature* **254**, 295 (1975).
4. Young, P. J., Shields, G. A. & Wheeler, J. C. *Astrophys. J.* **212**, 369 (1977).
5. Seidl, F. G. P. & Cameron, A. G. W. *Astr. Space Sci.* **15**, 44 (1972).
6. Carter, B. & Luminet, J. P. Preprint (Observatory of Meudon, 1981).

7. Carter, B. & Luminet, J. P. Preprint (Observatory of Meudon, 1982).
8. Lightman, A. P. & Shapiro, S. L. *Astrophys. J.* **211**, 244 (1977).
9. Lidskii, V. V. & Ozernoi, L. M. *Soviet Astr. Lett.* **5**, 16 (1979).
10. Fowler, W., Caughlan, G. R. & Zimmerman, B. A. *Rev. Astr. Astrophys.* **13**, 69 (1975).
11. Kumar, S. *Low Luminosity Stars*, 257 (Gordon & Breach, New York 1969).
12. Ezer, D. & Cameron, A. G. W. *Can. J. Phys.* **45**, 3461 (1967).
13. Wallace R. K. & Woosley, S. W. Preprint (Lick Observatory, 1981).

The rare earth elements in seawater

Henry Elderfield & Mervyn J. Greaves

Department of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK

The depth distributions of La, Ce, Nd, Sm, Eu, Gd, Dy, Er and Yb in the oceanic water column are used to evaluate the marine geochemical cycle of the rare earth elements and their application as water-mass tracers.

THE chemistry of the rare earth elements (REE) makes them particularly useful in studies of marine geochemistry. They are an extremely coherent group so that their relative abundances can be used to deduce their sources in sedimentary deposits¹. However, certain subtleties of this chemical similarity give additional important information. First, there is an ordered variation in the stability constants of REE complexes which means that the partition of rare earths between marine phases may fractionate the light relative to the heavy rare earths. Second, although the REE are predominantly trivalent there are exceptions: cerium can exist as Ce (IV) and europium as Eu (II), and these elements may fractionate from the 3+ REE as a function of redox potential. Third, some of the REE are weakly radioactive. In particular the decay of ¹⁴⁷Sm ($t_{1/2} = 1.06 \times 10^{11}$ yr) to ¹⁴³Nd coupled with the fractionation of Sm/Nd in the continental crust relative to the mantle means that the ¹⁴³Nd/¹⁴⁴Nd ratio may be used to constrain the sources of the REE.

For these reasons the REE (and, recently, Nd isotopes) have been used to assess the origins and depositional environments

of modern sediments and have been used as a suite of chemical indicators of oceanic change^{2–12}. The full exploitation of such applications requires a detailed understanding of the geochemical interactions of the REE in oceans and marine deposits. Our understanding of geochemical cycles in the oceans has increased dramatically as a result of recent measurements of the distributions of certain trace elements in the oceanic water column^{13–22}. In such cases, problems of sampling, contamination and analysis have been overcome to produce "oceanographically consistent"¹⁸ profiles which can be interpreted both in terms of known hydrographical and chemical features and also in terms of processes of element supply and removal such as organic matter regeneration, aeolian deposition, hydrothermal input, diffusion from bottom sediments and particle scavenging. No systematic study has yet been made of the distributions of the rare earths in the oceanic water column, the only measurements of the concentrations of the REE group in seawater being on isolated samples by Balashov²³, Goldberg²⁴ and Hogdahl²⁵ during the 1960s and by ourselves recently⁶, together with recent determinations^{7,8} of Nd and Sm. For this

Table 1 REE concentrations in North Atlantic ocean water (10^{-12} mol kg⁻¹)

Depth (m)	La	Ce	Nd	Sm	Eu	Gd	Dy	Er	Yb
0	36.7	66.3	34.3	6.01	0.615	5.59	5.00	3.63	3.15
100	13.0	16.8	12.8	2.67	0.644	3.41	4.78	4.07	3.55
200	17.0	22.3	15.8	4.52	0.849	—	5.31	4.62	4.07
600	22.5	18.4	19.7	3.86	0.796	4.85	5.41	4.58	4.14
700	25.2	24.7	21.9	4.23	0.757	5.23	5.43	4.57	4.07
900	20.8	9.64	21.1	4.32	0.823	5.20	5.61	4.94	4.66
1,000	—	20.8	22.8	4.51	1.01	—	6.00	—	—
1,500	22.8	9.71	19.0	3.72	0.954	5.31	6.03	5.30	4.99
2,500	29.4	26.1	25.0	4.75	0.895	7.19	6.10	5.09	4.79
3,000	32.6	19.3	25.4	4.69	0.987	5.80	6.14	5.33	5.21
4,500	54.4	55.1	45.8	8.25	1.22	8.27	6.83	5.34	5.16

reason we have determined the concentrations of nine rare earth elements in a vertical profile from the North Atlantic Ocean.

The seawater samples were collected in February 1979 during RRS *Shackleton* cruise 1B/79 from 28°01'N 25°59'W in the Cape-Maderia Abyssal Plain, water depth 5,250 m (also the site of GEOSECS station 115). The hydrography of the station (see also Fig. 4a) consists of a mixed layer of ~100 m, North Atlantic central water down to ~900 m, a diffuse lens of Mediterranean water with its core at ~1,200 m but influencing the potential temperature-salinity relationship to 2 km or more, below which is North Atlantic deep and bottom water. Our mass-spectrometric isotope dilution method for seawater analysis has been described previously⁹, although some modifications have been made. Our latest analytical scheme is, briefly: the 0.4- μ m filtered samples of ~50 l are acidified to pH 2 and the REE spikes added. After equilibration, the REE are concentrated by co-precipitation with ferric hydroxide, separated from the sea salts by cation exchange and from Ba by mixed solvent anion exchange. They are analysed as a group in the sequence

Eu-Yb-Ce-Sm-Nd-La-Er-Dy-Gd by selective ionization (VG Isomass 54E mass spectrometer) with on-line computing²⁶. Analytical precision and accuracy determined by running standard solutions is typically $\pm 1\%$ (2σ) or better with ratios of sample REE to reagent plus column blanks of $>10^3$.

REE distribution in the oceanic water column

The REE are present at extremely low concentrations in seawater (Table 1), ~ 10 – 70×10^{-12} mol kg⁻¹ for La–Nd, ~ 0.5 – 1×10^{-12} mol kg⁻¹ for Eu and ~ 3 – 8×10^{-12} mol kg⁻¹ for Sm–Yb. Despite these ultra-trace concentrations, factors such as contamination do not seem to have been an insurmountable problem (although some unascrivable scatter is present) and the results display non-random vertical distributions (Fig. 1).

The profiles of all the REE are characterized by: (1) a deep-water enrichment, all showing significant increases in concentration towards the bottom; (2) extrema in the mid-water column ~1,000 m; (3) low but non-zero concentrations at the

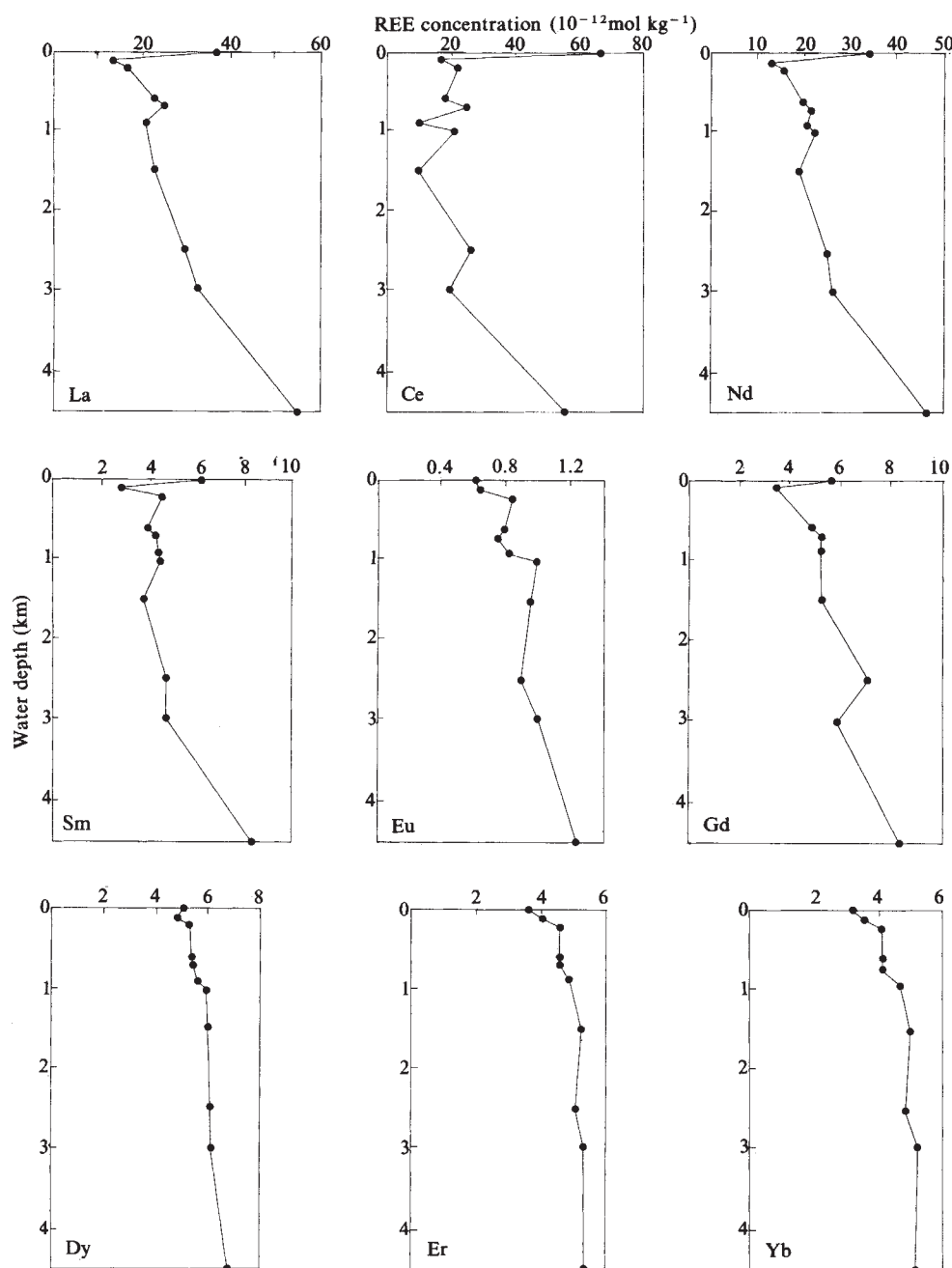


Fig. 1 Depth profiles of the REE at RRS *Shackleton* station 1B/79.

base of the mixed layer. In addition, the profiles of the light REE show surface maxima, that is, a surface-water enrichment.

The deep-water enrichment was suspected from earlier work¹. Some of the very early results are extremely high and must be regarded with suspicion. However, previous results^{7,8,24,25} show concentration levels similar to those reported here and include deep-water values which are higher than for shallow sample(s) from the same station. On the basis of this earlier work it has also been concluded¹ that REE levels in the deep Pacific are higher than those in the deep Atlantic. Modern data do not confirm this view. Indeed, a comparison of the most precise Nd data for the Atlantic (this study) and for the Pacific (the two analyses in refs^{7,8}) show that the Pacific values are ~30% lower at an equivalent water depth. Clearly, it is premature to make an assessment.

Table 1 also reveals that the deep-water enrichment varies in magnitude throughout the REE group in a systematic fashion. The proportional increase in concentration towards the bottom is greater for the light REE than for the heavy REE. For example, Yb-Dy are enriched by a factor of ~1.4 at 4.5-km depth relative to the base of the mixed layer (100 m) whereas Nd is enriched 3.6 times and La, 4.2 times. Like this deep-water behaviour, the surface-water enrichments of the REE vary with atomic number but with an important difference: the light REE are enriched (for example, La by a factor of 2.8 compared with the base of the mixed layer) whereas the heavy REE are not enriched or are slightly depleted. Because of their similarity, the surface-water and deep-water enrichments show a striking correlation (Fig. 2). Note that the enrichments of Ce and Eu are anomalous compared with values that may be predicted from interpolations between their respective REE neighbours. Both the fractionation of the light- from heavy-REE (or vice versa) at the surface and bottom boundaries of the water column and the anomalous behaviours of Ce and Eu are compatible with the chemistry of the REE group.

REE patterns

The shale-normalized REE pattern of seawater has been shown²⁴ to be characterized by a monotonic increase in the concentrations of the heavy REE relative to the light REE and by a marked depletion in Ce relative to its REE neighbours (that is, a negative cerium anomaly). The REE patterns obtained here may be placed in three groups (Fig. 3):

(1) The pattern for surface seawater is unlike any reported previously. The pattern is flat and shale-like with a slight negative Ce anomaly, but with a prominent negative Eu anomaly.

(2) Seawater below the mixed layer has a pattern similar to that reported previously although with a less striking Ce anomaly. With increasing depth the heavy REE enrichment becomes less significant, a negative Eu anomaly develops and the negative Ce anomaly strengthens and then weakens (Fig. 4). As a consequence, the pattern for the deepest sample

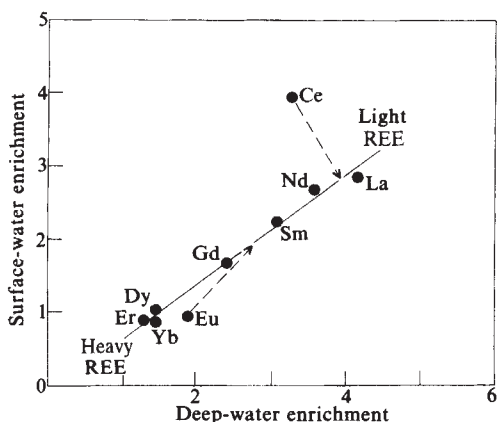


Fig. 2 Surface-water/deep-water enrichments of the REE relative to concentrations at base of the mixed layer.

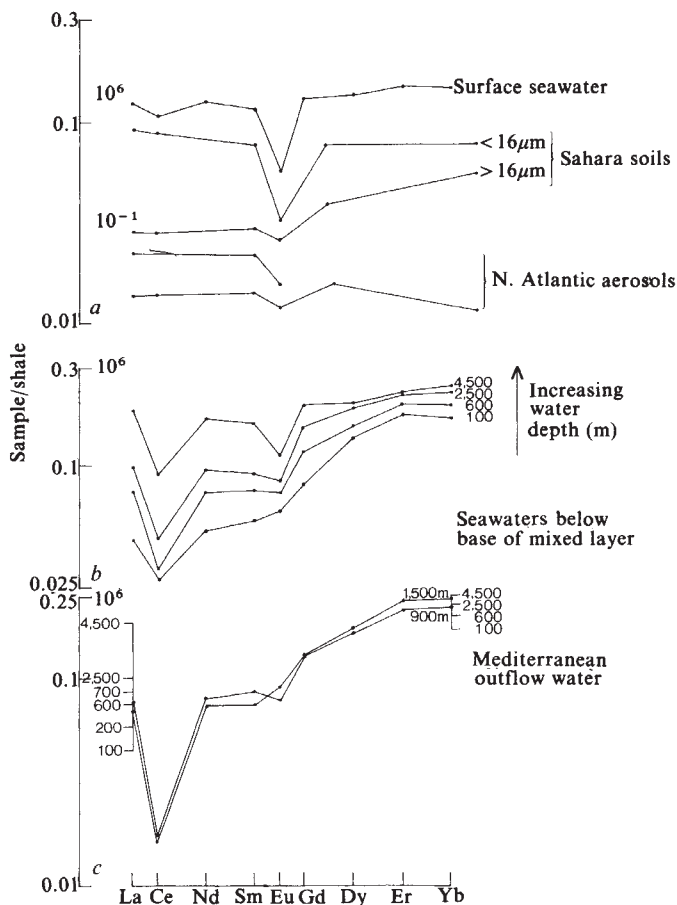


Fig. 3 REE patterns: a, surface seawater compared with Sahara soils and North Atlantic aerosols; b, seawaters below base of the mixed layer, 100–4,500 m; c, seawaters in the core of Mediterranean water, 900–1,500 m.

analysed begins to resemble that for surface seawater except for its Ce anomaly.

(3) The seawaters between ~900 and 1,500 m, representing Mediterranean overflow water, also show 'typical' seawater patterns but are anomalous relative to waters above and below in displaying stronger negative Ce anomalies, weaker negative Eu anomalies, and higher heavy REE and lower light REE concentrations (Fig. 3c) give a more marked heavy REE enrichment (Fig. 4).

Boundary layer processes

Because of the deep-water enrichment of the REE and the contention that REE are higher in the Pacific than the Atlantic, the REE have been likened to nitrate and phosphate and to those trace elements whose depletion in surface waters has been interpreted in terms of biological removal^{1,7,8}. From the present results it seems unlikely that the REE exhibit simple nutrient-like behaviour in seawater. Overall, the REE profiles are dissimilar to the profiles both of phosphate (which reflects near-surface regeneration of soft tissues) and of silicate (which reflects deep-water regeneration of hard skeletal materials) except that all increase with depth (Figs 1, 4). Therefore, one hypothesis is that the behaviour of the REE in waters below the mixed layer is nutrient-like, the light REE following Si and the heavy REE following P. However, the surface-water behaviour of the REE is not nutrient-like. Both silicate and phosphate show near-zero concentrations in surface waters unlike the REE. The light REE also show a significant surface maximum.

The profiles for the light REE are reminiscent of those of Cu (ref. 18) Al (ref. 21) which have been interpreted in terms of inputs across the surface and bottom boundaries of the oceans coupled with scavenging by a combination of biological removal

and general (inorganic?) removal with subsequent incorporation in and removal by zooplanktonic faecal matter or uptake on surfaces of skeletal fragments. For the REE, the association with colloidal matter such as iron-oxide flocs inside tests, or as coatings, or in biogenic particles has been favoured¹⁰. Although we cannot yet identify any specific carrier for the REE we can show from the general depth trends of the REE and of P, Si and alkalinity (Table 2a) that (1) planktonic material contains insufficient REE to account for the observed deep-water enrichments of the light REE (little data are available for the heavy REE); (2) REE in foraminiferal carbonate can just account for the heavy REE enrichments but not for the light REE; (3) diatom opal can readily account for all the enrichments. REE/P/Si/Ca ratios in large particles from sediment traps suggest that such material contains a large excess of REE compared with the levels needed to satisfy the deep-water enrichments. Hence, we propose that the REE are removed selectively from solution in surface waters by scavenging, the light REE being removed more effectively than the heavy REE. This process is consistent with the previously observed REE pattern for seawater which led Goldberg²⁴ to suggest that the heavy REE enrichment may be attributed either to the increasing stability of complexes of the heavier REE with seawater ligands or to their differential adsorption on solid phases. Our observations clearly indicate that Ce behaves differently from the other REE. The fractionation of the light and heavy 3+ REE is at its maximum at the base of the mixed layer (the same is true for the least negative Eu anomaly) showing that the surface removal processes are restricted to this zone (Fig. 4). In contrast, the negative Ce anomaly develops over a depth of 500 m or more, pointing to a different process, presumably the oxidation of Ce³⁺ to the less soluble 4+ valency state.

As well as REE removal in surface waters there is tentative evidence for deep-water removal. Below ~2 km where the potential temperature-salinity relationship is linear, plots of the REE against either of these conservative tracers show negative curvature, indicating removal of the REE in deep waters. However, the deep circulation of the Atlantic is complex and deep-water scavenging is not necessarily implied by this observation.

What are the causes of inputs at the surface and bottom boundary layers? Analogy with Cu and Al suggests that the bottom source is the interstitial waters of marine sediments into which the REE are released during early diagenesis when the REE-containing phases dissolve or lose their binding capacities. Again, the difficulty of modelling chemical gradients in Atlantic deep waters means that no other benthic flux need be postulated other than from regeneration. However, a sediment source is consistent with our work on oceanic ferromanganese nodules and sediments which indicates that the REE are mobile during early diagenesis^{6,10,11}. The diffusive benthic fluxes which would

be required to maintain this deep-water enrichment are considerable—there must be strong REE gradients in the interstitial waters. Minimum benthic fluxes calculated using residence times of 100–400 yr for North Atlantic deep water are comparable in magnitude to the average rates of REE accumulation in deep-sea sediments^{1,27}. Hence, if diffusion from sediments is an important source, the uppermost layer of deep sea sediments should be strongly enriched in REE, as indeed seems to be the case for La (ref. 28). Because the cycling of the REE through the surface layer of sediments seems to be such an important part of the oceanic REE cycle it is obvious why ferromanganese nodules and coexisting sediments have the same Nd isotopic composition as deep ocean water^{7,8,10,12}. Nodules obtain Nd through sediment diagenesis which also provides the Nd in bottom waters which itself is scavenged by sedimenting particles.

The source of the surface-water enrichment relative to the bottom of the mixed layer in light REE is probably aeolian. The other trace elements for which surface-water enrichment has been observed are Cu, Mn, Pb and Al (refs 18–21). Analogy with atmospheric input tracers ²¹⁰Pb and ⁷Be which also show surface enrichment suggests that the inputs are aeolian. For the REE it is possible to employ the 'internal' tracer of the REE pattern for comparison with potential input material. This exercise (Fig. 3a) shows that the striking REE pattern for surface seawater at our station is comparable with that for Sahara Desert soils and for the North Atlantic (Sahara) marine aerosol²⁹. Nd isotope studies on the water and aerosols would be useful to confirm the interpretation but atmospheric input is probably the major source of the REE in North Atlantic surface water in this region.

Hydrographical signals

We cannot rule out the possibility that certain parts of the vertical profiles may be primarily hydrographical signals. First, the surface values could reflect advected input of continental REE and the deep-water values could mean that North Atlantic deep and bottom water forms with very high REE contents. Second, and more plausible, the mid-water extrema in the REE profiles probably reflect a distinctly different composition of Mediterranean water compared with the Atlantic water masses. Clearly, water in the Mediterranean outflow, centred at ~1,200 m at this station (Fig. 4), has a distinctive REE pattern compared with adjacent waters (Fig. 3): that is, a larger heavy REE enrichment, a more negative Ce anomaly and a less negative Eu anomaly (Fig. 4). We found the ¹⁴³Nd/¹⁴⁴Nd ratio of biogenic carbonate from the Mediterranean to be 0.512275 ± 20 , slightly higher than for Atlantic ratios^{7,8}. Hence, the use of REE patterns and Nd isotopic ratios as modern and palaeo-oceanographic tracers of water masses and ocean basins seems feasible.

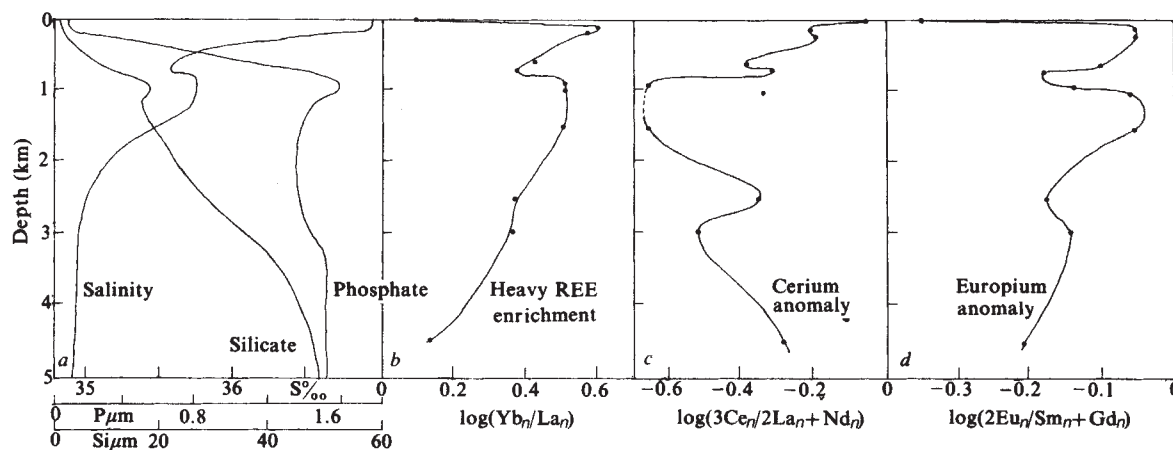


Fig. 4 Depth profiles for hydrographic and nutrient features for station 1B/79 compared with profiles of shale-normalized REE parameters: a, salinity, phosphate and silicate; b, heavy REE enrichment; c, cerium anomaly; d, europium anomaly.

Table 2 Enrichment ratios and residence times of the REE

a, Comparison of REE enrichment ratios for possible carriers with elemental compositions									
	La	Ce	Nd	Sm	Eu	Gd	Dy	Er	Yb
Plankton									
$\Delta\text{REE}/\Delta\text{P}$ in water column (mol ratio $\times 10^6$)	28	26	22	3.8	0.39	3.3	1.4	1.2	1.4
REE content calculated from ratio*/measured content†	19	10	70	—	—	—	19	—	—
10^3 ratio/(REE/P) in sediment trap and faecal materials‡	3–19								
Opal									
$\Delta\text{REE}/\Delta\text{Si}$ in water column (mol ratio $\times 10^6$)	0.86	0.79	0.68	0.12	0.012	0.10	0.042	0.035	0.042
REE content calculated from ratio§/measured content†	0.14	0.090	0.098	0.077	0.030	0.062	0.027	0.039	0.056
10^3 ratio/(REE/Si) in sediment trap and faecal materials‡	18–21								
Carbonate									
$\Delta\text{REE}/0.5\Delta\text{alkali}$ in water column (mol ratio $\times 10^6$)	4.1	3.8	3.3	0.56	0.058	0.49	0.21	0.17	0.20
REE content calculated from ratio /measured content†	4.5	15	4.8	4.2	1.5	2.7	1.0	1.0	1.2
10^3 ratio/(REE/Ca) in sediment trap and faecal materials‡	22–164								
b, Residence times of the REE									
Water column standing crop ($10^{-9}\text{mol cm}^{-2}$)¶	14	13	13	2.4	0.44	2.8	2.7	2.3	2.1
Residence time in surface ocean relative to aeolian input (yr)#									
Maximum	6.2	4.4	6.6	7.8	19	11	17	20	20
Minimum	0.97	0.87	1.0	0.93	1.1	1.1	1.2	1.3	1.2
Overall oceanic residence time (yr)									
Relative to river input**	1,500	720	1,100	1,100	910	1,200	1,400	2,000	2,200
From aeolian + river input††	240	110	450	220	470	300	340	420	410
From sediment output‡‡	690	400	620	600	460	720	810	980	970

* Assuming molecular weight of plankton = 1,527 ($\text{C}_{106}\text{N}_{16}\text{P}_1$); † from ref. 10; ‡ from ref. 37; § assuming (as in ref. 22) molecular weight of opal = 150.1 ($\text{SiO}_2 \cdot 5\text{H}_2\text{O}$); || assuming $\Delta\text{Ca}^{2+} = 0.5\Delta$ alkalinity; ¶ integrated areas in Fig. 1; # see text for methods of calculation; ** standing crop $\times$ oceanic area of $361 \times 10^{16} \text{ cm}^2$ /REE in river water (based on ref. 9) $\times$ global river discharge rate of $4 \times 10^{19} \text{ cm}^3 \text{ yr}^{-1}$. Values not corrected for REE loss in estuaries which would increase τ by factor of ~ 4 ; †† assumes La flux of $60 \times 10^{-12} \text{ mol cm}^{-2} \text{ yr}^{-1}$ (50 aeolian—see text; 10 rivers) and other REE pro rata. Probably minimum values of τ because of relatively large aeolian flux in north-east Atlantic; ‡‡ using Piper's estimates (ref. 1) of average REE accumulation rates in sediments.

The oceanic REE cycle

From the above discussion we can summarize the geochemical cycle of the REE at our North Atlantic station and compute their oceanic residence times. As a first approximation we will assume a steady state and ignore horizontal transport. The major surface input appears to be atmospheric and the REE are leached from aeolian particles without significant fractionation (except possibly for Ce and Eu). The dissolved REE are scavenged from surface waters with preferential removal of the light REE. Thus, the residence time (τ) of La in the surface ocean must be less than that of Yb, the upper limit for the least reactive heavy REE being the residence time of the mixed layer ($\tau_{\text{H}_2\text{O}}$) of ~ 20 yr (ref. 30). For typical particulate loadings in surface seawater near station 1B/79 of $\sim 30 \mu\text{g per kg}$ (refs 31, 32) and for the enrichment of double crustal abundance of the REE in marine aerosols²⁹, there would be $\sim 7.5 \times 10^{-12} \text{ mol kg}^{-1}$ of particulate La in surface seawater if no leaching occurs at all, compared with $\sim 37 \times 10^{-12} \text{ mol kg}^{-1}$ of dissolved La. Therefore τ_{La} in surface water must be $37/7.5$ times the residence time of unreactive particles ($\sim 0.2 \text{ yr}^{33}$) that is ~ 1 yr, if all the REE are leached from the aeolian input. This is a minimum value for the residence time because if, say, only 15% dissolves then $\tau_{\text{La}} \sim 6$ yr. Because τ_{La} must be $< \tau_{\text{H}_2\text{O}}$ then at least $\sim 5\%$ of the REE must be leached.

Another approach to estimating τ is from the REE fractionation below, relative to in, the mixed layer. As the Yb/La ratio decreases by a factor of 3.2, the upper limit for τ_{La} in the mixed layer is about one-third of $\tau_{\text{H}_2\text{O}}$, that is 6.2 yr. Note the anomalous positions of Ce and Eu in the residence times computed in these ways (Table 2b). τ_{Ce} is shorter and τ_{Eu} is longer than their position in the REE group might imply.

These estimates for surface ocean residence times for the REE are compatible with aerosol data. From τ and the amount of a rare earth element present in the mixed layer (a) the input flux can be predicted from a/τ . The figure for La is $58 \times 10^{-12} \text{ mol cm}^{-2} \text{ yr}^{-1}$ (using τ_{max}). In comparison, typical Al deposition rates^{33,34} of $\sim 5 \mu\text{g cm}^{-2} \text{ yr}^{-1}$ together with La/Al in aerosols which average 14×10^{-4} by weight (using data in ref. 29; La and Al seem to correlate in marine particulates) gives a similar flux of $50 \times 10^{-12} \text{ mol cm}^{-2} \text{ yr}^{-1}$. Alternatively, La in the North Atlantic aerosol has been measured^{34,35} at 0.28 ng m^{-3} (using typical Al in aerosols^{33–35} of 150 ng m^{-3} and the above La/Al ratio gives a similar value, 0.21 ng m^{-3}). Using a typical depositional velocity³⁴ of 1 cm s^{-1} gives $64 \times 10^{-12} \text{ mol cm}^{-2} \text{ yr}^{-1}$.

The rapid transport to the deep ocean of the REE scavenged in the surface ocean and their release from the sediment explains the similar REE patterns of deep and surface waters and the link in their enrichments shown in Fig. 2. Thus, the surface-water enrichments relative to 100 m depth (Fig. 1) of the light REE reflects their preferential removal by scavenging, the sedimented products of which release the REE during early diagenesis leading to stronger deep-water gradients for the light than for the heavy REE. However, we cannot yet rule out *in situ* regeneration as a source of the REE in deep waters. The Ce anomaly in Fig. 2 might be expected from its redox chemistry in surface seawaters and sediments. In surface waters, Ce^{3+} released from particles will oxidize and so be removed relative to the 3+ REE. In deep-sea sediments this same reaction means that Ce released from decomposing particles will be less mobile than the 3+ REE (this is observed in terms of the positive Ce anomaly in most oceanic ferromanganese nodules¹) leading to a relatively weaker deep-water gradient for Ce.

The surface and deep sources for the REE suggested here can explain the imbalance in the overall REE budget recognized by Piper¹ and Martin³⁶ who attempted to balance sediment removal with river input only. The residence times with respect to river input (Table 2b—these are minimum values as they are uncorrected for REE removal which occurs in at least some estuaries^{9,36}) are much too long to represent the total input in view of the water mass effects described earlier. A major atmospheric input will redress this balance (Table 2b). Overall, accurate values for τ are difficult to estimate because the atmospheric fluxes into the north-east Atlantic are probably much larger than the oceanic average. However, the atmospheric fluxes may be highly significant. Equally, a benthic flux of the magnitude of the river flux would allow the long riverine residence times to be compatible with the short scavenging residence times. Thus, overall, the REE may reside in the ocean for ~500 yr before being scavenged but are regenerated on average, say, five times before being finally buried and therefore would have residence times with respect to river input of

2,500 yr. The computed values for the overall residence times of the REE (Table 2b) being less than the oceanic mixing time are compatible with the tracing of water masses using REE contents as suggested by the Mediterranean water. In particular, τ for the light REE are short. Thus, the difference between water masses should be more pronounced for the light REE than for the heavy REE, as is observed (Fig. 2). In addition, the value for τ_{Nd} is consistent with the inter-oceanic differences in the $^{143}Nd/^{144}Nd$ ratio of seawater^{7,8,12}. Although the above described residence times have various shortcomings and uncertainties due to the limited data base and inherent assumptions, their trends clearly reflect our suggested input and removal mechanisms and our basic conclusion that the REE are rapidly cycled in the marine environment.

We thank members of RRS *Shackleton* cruise 1B/79 and oceanographic and isotopic colleagues at Leeds for their help; and W. & T. Avery Ltd for the use of scales to weigh the ~1,000 kg of seawater. This research was supported by NERC grants GR3/3125 and GR3/4444.

Received 21 October 1981; accepted 21 January 1982.

- Piper, D. Z. *Chem. Geol.* **14**, 285 (1974).
- Wildeman, T. R. & Haskin, L. A. *Geochim. cosmochim. Acta* **37**, 419 (1973).
- Bender, M. L. *et al. Earth planet Sci. Lett.* **12**, 425 (1971).
- Robertson, A. H. F. & Fleet, A. J. *Earth planet Sci. Lett.* **28**, 285 (1976).
- Glasby, G. P. *Mar. Chem.* **1**, 105, 1973).
- Elderfield, H. & Greaves, M. J. *Earth planet. Sci. Lett.* **55**, 163 (1981).
- Piepgas, D. J., Wasserburg, G. J. & Dasch, E. J. *Earth planet. Sci. Lett.* **45**, 223 (1979).
- Piepgas, D. J. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **50**, 128 (1980).
- Elderfield, H. & Greaves, M. J. in *Trace Metals in Sea Water* (ed. Wong, C. S.) (Plenum, New York, in the press).
- Elderfield, H., Hawkesworth, C. J., Greaves, M. J. & Calvert, S. E. *Geochim. cosmochim. Acta* **45**, 513 (1981).
- Elderfield, H., Hawkesworth, C. J., Greaves, M. J. & Calvert, S. E. *Geochim. cosmochim. Acta* **45**, 1231 (1981).
- Goldstein, S. L. & O'Nions, R. K. *Nature* **292**, 324 (1981).
- Boyle, E. A., Sclater, F. & Edmond, J. M. *Nature* **263**, 42 (1976).
- Sclater, F., Boyle, E. & Edmond, J. M. *Earth planet. Sci. Lett.* **31**, 119 (1976).
- Bender, M. L. & Gagner, C. L. *J. mar. Res.* **34**, 327 (1976).
- Bruland, K. W. *Earth planet Sci. Lett.* **47**, (1980).
- Measures, C. I. & Burton, J. D. *Earth planet. Sci. Lett.* **46**, 385 (1980).
- Boyle, E. A., Sclater, F. R. & Edmond, J. M. *Earth planet Sci. Lett.* **37**, 38 (1977).

- Klinkhammer, G. P. & Bender, M. L. *Earth planet. Sci. Lett.* **46**, 361 (1980).
- Schaule, B. K. & Patterson, C. C. *Earth planet. Sci. Lett.* **54**, 97 (1981).
- Hydes, D. J. *Science* **205**, 1260 (1979).
- Froelich, P. N. & Andrae, M. O. *Science* **213**, 205 (1981).
- Balashov, Y. A. & Khitrov, L. M. *Geochim. Int.* **9**, 877 (1961).
- Goldberg, E. D., Koide, M., Schmitt, R. A. & Smith, R. H. *J. geophys. Res.* **68**, 4209 (1963).
- Hogdahl, O. T., Melson, S. & Bowen, V. *Adv. Chem. Ser.* **73**, 308 (1968).
- Thirlwall, M. *Chem. Geol.* (in the press).
- Brewer, P. G., Nozaki, Y., Spencer, D. W. & Fleer, A. P. *J. mar. Res.* **38**, 703 (1980).
- Bonatti, E., Fisher, D. E., Joensuu, U. O. & Rydell, H. S. *Geochim. cosmochim. Acta* **35**, 189 (1971).
- Rahn, K. A. *The Chemical Composition of the Atmospheric Aerosol* (University of Rhode Island, Tech. Rep., 1976).
- Montgomery, R. B. in *The Atmosphere and the Sea in Motion* (ed. Bolin, B.) (Rockerfeller, New York, 1958).
- Brewer, P. G. *et al. Earth planet Sci. Lett.* **32**, 393 (1976).
- Krishnaswami, S., Lal, D. & Somayajulu, B. L. K. *Earth planet. Sci. Lett.* **32**, 403 (1976).
- Wallace, G. T., Hoffman, G. L. & Duce, R. A. *Mar. Chem.* **5**, 143 (1977).
- Buat-Menard, P. & Chesselet, R. *Earth planet. Sci. Lett.* **42**, 399 (1979).
- Dudley, N. D., Ross, L. E. & Nashkin, V. E. in *Modern Trends in Activation Analysis* Vol. 1 (eds DeVoe, J. R. & LeFleur, P. D.) (NBS Spec. Publ. 312, 1969).
- Martin, J.-M., Hogdahl, O. & Philippot, J. C. *J. geophys. Res.* **81**, 3119 (1976).
- Spencer, D. W. *et al. J. mar. Res.* **36**, 493 (1978).

Genome instability in a region of human DNA enriched in *Alu* repeat sequences

Bruno Calabretta*, Donald L. Robberson, Hugo A. Barrera-Saldaña*,
Thymios P. Lambrou† & Grady F. Saunders*

Departments of Biochemistry*, Molecular Biology, and Developmental Therapeutics†, University of Texas System Cancer Center, Houston, Texas 77030, USA

University of Texas Graduate School of Biomedical Sciences, Houston, Texas 77030, USA

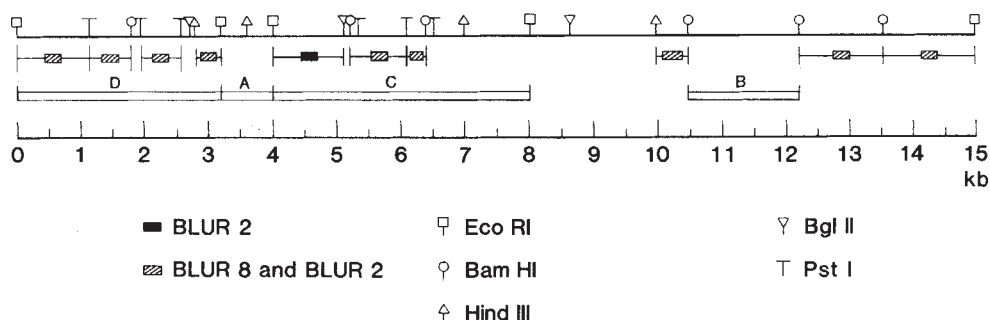
DNA rearrangements involving restriction fragment length polymorphism and variation in copy number were detected in the human genome by blot hybridization with a cloned segment of human DNA initially present in a cluster of Alu repeat sequences. These rearrangements involve both extrachromosomal circular duplex DNAs and integrated sequences, indicating the presence of transposable elements in human cells.

RENATURATION kinetics¹ and genomic library screening² have shown that the human genome contains at least 400,000 copies of the 300-base pair (bp) *Alu* repeat sequence, the predominant repetitive sequence in man^{3,4}. Members of the *Alu* repeat family have been proposed to have cellular functions on the basis of various characteristics, including the widely dispersed arrangement of *Alu* repeats, the conservation of a 40-bp sequence of the *Alu* repeat in mammalian evolution, the homology of a 14-bp segment with a sequence near the replication origin in several DNA viruses, and the observation that the *Alu* repeat is represented in transcripts generated by both

RNA polymerase II and III^{5,6}. Thus, a role for *Alu* repeat sequences in the initiation of DNA replication and nuclear RNA processing has been postulated^{5,6}.

The most striking feature of the *Alu* repeat family is its large numerical representation in the human genome, which suggests that *Alu* repeat sequences might be involved in genetic rearrangements, a role which could be identified if we consider the human genome to be a dynamic structure. Although most members of the *Alu* family are scattered throughout the human genome, some may be clustered in certain genomic regions. Such an arrangement would provide a good opportunity to test

Fig. 1 Localization of *Alu* repeat sequences within the restriction enzyme cleavage map of λ H15. λ H15 was isolated from a human genomic library⁷ as a clone hybridizing to total nick-translated repetitive sequences, grown as previously described², and the human insert was characterized with respect to the distribution of repetitive sequences. The orientation of the human insert with respect to the left and right arms of the vector was facilitated by knowing the restriction map for Charon 4A²⁴. The localization of *Alu* repeat sequences was initially derived for the restriction map of λ H15, and confirmed for the restriction map of several segments of λ H15 (subcloned in to pBR322 or pACYC184 and indicated by segments D, A, C and B, respectively) using Southern blot hybridization²⁵ with nick-translated *Alu* repeat inserts from either the clone BLUR 8 or BLUR 2 (ref. 4) as previously described⁶.



the hypothesis that repetitive sequences facilitate genetic rearrangements.

Alu repeat sequences are clustered in regions of the human genome

We searched a human genomic DNA library⁷ for clones containing several *Alu* repeat sequences, having previously shown that more than 95% of this library contained members of the *Alu* repeat family² and that in several clones, two or more copies were present (unpublished observation). Clone λ H15, randomly selected as strongly hybridizing to repetitive sequences, was examined in detail. An initial digestion of this clone with several restriction enzymes, followed by Southern blot hybridization with the cloned *Alu* repeat sequence probes BLUR 2 and BLUR 8 (ref. 4), gave several hybridizing bands, indicating that many *Alu* repeat family members were present in this 15-kilobase (kb) DNA insert. Double digestion with restriction

endonucleases *EcoRI* plus *BamHI* and *EcoRI* plus *PstI* showed 7 and 12 bands, respectively, hybridizing to these cloned members of the *Alu* repeat family.

Figure 1 shows the restriction enzyme cleavage map of λ H15, together with the approximate locations of the *Alu* repeat family members. At least 10 *Alu* repeat sequences are clustered in this clone, a minimal estimate due to the high degree of divergence among different members of the *Alu* repeat family⁴. Two observations support this assertion. First, we identified an *Alu* repeat sequence hybridizing very strongly to BLUR 2 (dark box in Fig. 1) and very weakly to BLUR 8. The sequence divergence between BLUR 8 and BLUR 2 is 24%, thus precluding cross-hybridization in the conditions used. This observation suggests that some members of the *Alu* repeat family have diverged considerably in the *Alu* repeat cluster of λ H15. Second, in R-loop hybridization between poly(A)⁺ RNA, derived from a human lymphoblastoid line and subclone pAH15C (Fig. 1), two additional *Alu*-like sequences have been detected and positioned with respect to the *Alu* sequence distribution derived by Southern blot analysis⁶. We may therefore consider the segment of λ H15 between the first and the fourth *EcoRI* sites (Fig. 1) as a cluster of *Alu* repeat sequences.

A role for *Alu* repeat sequences in genomic rearrangements

The unusual structure of this region provides an opportunity to test the hypothesis that *Alu* repeat clustering may be involved in genome rearrangements. To investigate this possibility we took advantage of the observation that in λ H15, DNA segments of single copy or very low repetition frequency are flanked on both sides by *Alu* repeats. Subclones of two of these regions (A and B in Fig. 1) were constructed for use as hybridization probes to detect restriction fragment length polymorphisms (RFLPs) in the genomic region in which the human insert of λ H15 is localized. If DNA rearrangements occur in this region at a relatively high frequency, they should be detectable by RFLP analysis. We therefore used cleavage with restriction enzymes and blot hybridization experiments to examine the DNA from tissues of healthy individuals and patients with neoplastic haematological diseases, and found substantial polymorphism with DNA from tissues of healthy individuals using both A and B region probes (Fig. 2). Digestion of several of the DNA samples with other restriction enzymes also revealed patterns of polymorphism. Furthermore, we found that the intensity of one hybridizing band differed markedly among DNAs derived from leukocytes of a normal donor (Fig. 2, lane c, left side) and DNA derived from leukaemic leukocytes (Fig. 2, lanes d-i). These results indicate both variation in the number of hybridizing sequences (copy number) and multiple forms revealed by RFLP of the sequence in human DNA that hybridizes to pAH15A.

Using a gene copy number reconstitution experiment, we determined and compared the number of sequences hybridizing to pAH15A in DNAs derived from leukocytes of a patient with acute lymphocytic leukaemia and a normal individual. The method used has sufficient sensitivity and reproducibility to

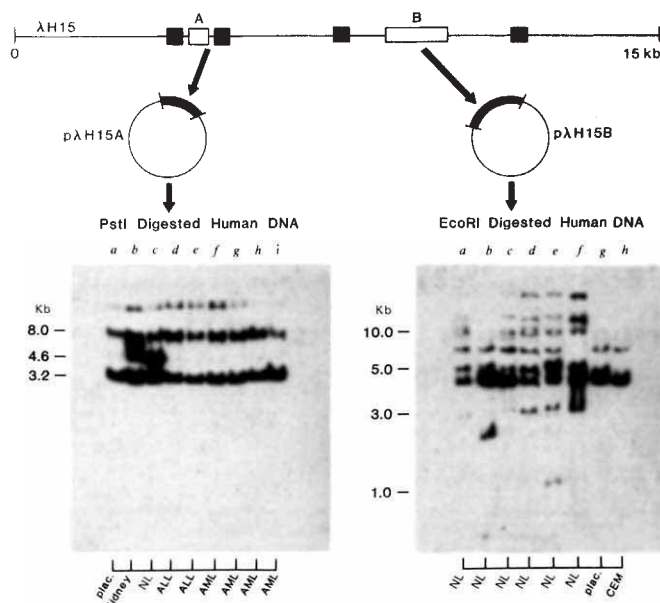
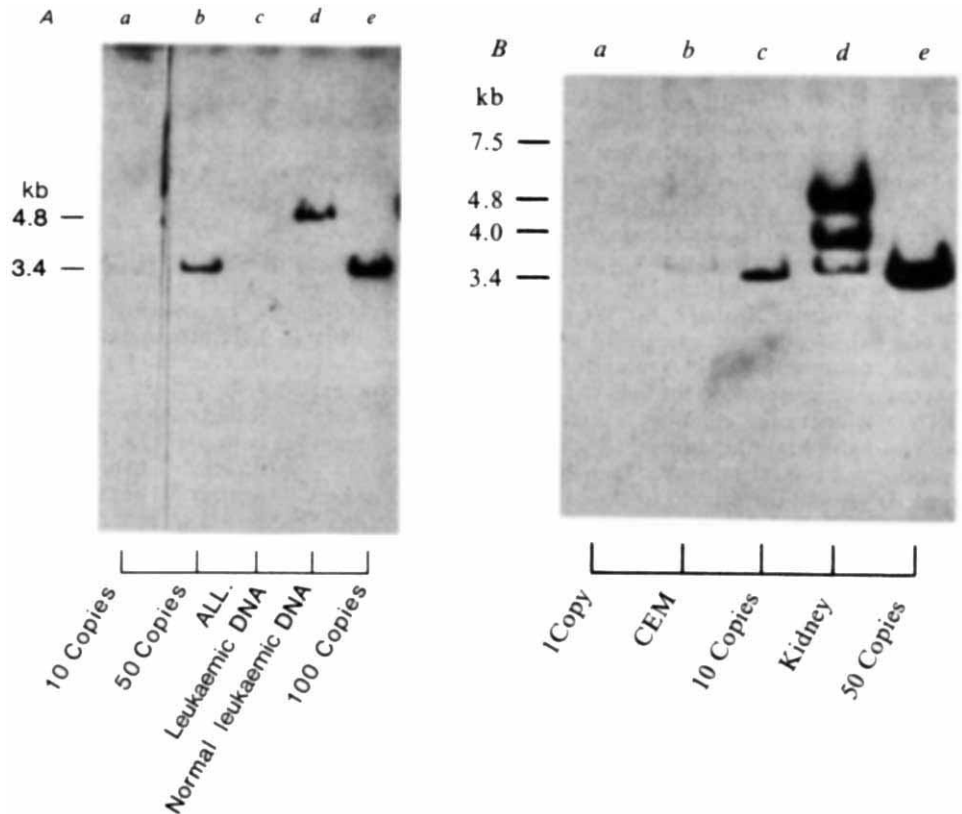


Fig. 2 Experimental design used to investigate DNA rearrangements in tissues of normal and leukaemic individuals. The dark boxes on the human insert of λ H15 represent *Alu* repeat sequences flanking the DNA fragments subcloned in pACYC184 (pAH15A) and pBR322 (pAH15B). Subclones were prepared as previously described⁶. Human DNA (10 μ g) prepared as described elsewhere²⁶ was digested with an excess of *PstI* (left side) or *EcoRI* (right side) for 2 h and electrophoresed through 1% agarose gels in 50 mM Tris (pH 8.2), 50 mM boric acid, 10 mM EDTA at 80 V for 6 h. After electrophoresis, the DNA was transferred to nitrocellulose by the method of Southern²⁵ and the resulting blots were hybridized with 2×10^7 c.p.m. of labelled pAH15A or pAH15B DNA, prepared by nick translation²⁷. The hybridization was performed in the presence of 10% dextran sulphate²⁸ as previously described²⁹. Filters were exposed for 12 h to X-ray films with intensifier screens. The approximate sizes of DNA in kilobase pairs (kb) in the bands are indicated on the left. plac, Placenta; NL, normal leukocytes; AML, acute myelogenous leukaemia; ALL, acute lymphocytic leukaemia; CEM, a human lymphoblastoid cell line. Each lane contains DNA from a different individual.

Fig. 3 A, determination of the number of human genomic DNA sequences complementary to p λ H15A. The number of sequences hybridizing to p λ H15A was determined as follows: 5 μ g of DNA derived from leukaemic lymphocytes of a patient with acute lymphocytic leukaemia (ALL) (lane c) and from peripheral white blood cells of a normal individual (lane d) were digested with *Bam*HI and electrophoresed on a 1% agarose gel with various amounts of *Bam*HI-digested λ H15 DNA from which p λ H15A was derived. The copy number reference, *Bam*HI-digested λ H15, was applied in amounts corresponding to 10, 50 and 100 copies per haploid genome in 5 μ g of human DNA (lanes a, b and e respectively). These lanes also contained 5 μ g of *Eco*RI-digested chicken DNA used as carrier. Assuming a human haploid genome size of 3×10^9 bp, the 45-kb λ H15 constitutes 1.53×10^{-5} of the human haploid genome. Therefore, the amount of λ H15 in 5 μ g of human DNA, if present at one copy per haploid genome, is 7.6×10^{-5} μ g. To apply 10, 50 and 100 copies, 4.5 μ g of λ H15 DNA were digested with *Bam*HI, and the reaction was stopped by adding one-fifth volume of a solution containing 50% glycerol, 1% SDS and 0.5% bromophenol blue. Appropriate dilutions were made from this solution. B, number of genomic sequences complementary to p λ H15A in DNA of a human lymphoblastoid cell line (CEM) and normal kidney. DNA isolated from CEM cells and kidney was electrophoresed with 1, 10 and 50 copies of *Bam*HI-digested λ H15 DNA used as an internal standard, and then blot-hybridized to p λ H15A as described earlier. Densitometric scans of the autoradiograms demonstrate that the band at 3.4 kb hybridizing to p λ H15A in CEM cells corresponds to three copies in CEM cells and to approximately six copies in kidney tissue.



give reliable estimates of gene copy numbers^{8,9,10}. Titration experiments (Fig. 3A) show that more than 50 copies of sequences hybridizing to p λ H15A are present in the leukocytes of the normal individual, compared with less than 5 copies in leukaemic leukocytes. In all the patients with acute leukaemia so far examined (10 cases) we found less than five copies of p λ H15A-hybridizing sequences. In contrast, of six patients with chronic myelogenous leukaemia examined, four showed an autoradiographic density comparable with that of healthy individuals, whereas two had copy numbers comparable with that found in acute lymphocytic leukaemia.

The number of sequences hybridizing to p λ H15A in DNA derived from leukocytes of normal donors was rather variable but was usually more than 50 copies per haploid genome equivalent. In control experiments, the DNAs derived from normal and leukaemic leukocytes were hybridized with a recombinant plasmid containing the coding region of the human placental lactogen (HPL) gene, present in three to four copies per haploid genome¹⁰, and equal intensities of hybridizing bands were observed in both DNA groups. We can therefore exclude variations in hybridization efficiency as an explanation of the differences in copy number observed using the p λ H15A DNA as probe.

Figure 3B shows an additional example of variations in copy number of sequences hybridizing to p λ H15A. A gene copy number reconstitution experiment reveals that in DNA derived from either a normal kidney tissue or a human lymphoblastoid cell line, a meaningful variation exists in the content of sequences hybridizing to p λ H15A. The bands at 4.8 and 4.0 kb observed in kidney DNA (Fig. 3B, lane d) are not detectable at all in CEM DNA (Fig. 3B, lane b). The only band in common with these DNAs has a size of 3.4 kb and is present at three copies in the lymphoblastoid cell line and six copies in the kidney tissue per haploid genome. The 3.4-kb band, common to the human DNAs shown in Fig. 3, corresponds to the region of clone λ H15 between the *Bam*HI sites at 1.8 and 5.2 kb shown in Fig. 1. We therefore identified variations in copy

number of a specific genomic sequence of human DNA and the results described below indicate that some of the hybridizing bands correspond to circular extrachromosomal molecules.

DNA rearrangements exist in different tissues of the same individuals

Having found variations in copy number and RFLPs in this region of the genome, we next investigated whether different forms of the hybridizing sequence occurred within different tissues of the same individual. A surprising result was that when DNA isolated from a 22-yr old woman who had died of acute lymphocytic leukaemia was digested with *Eco*RI and hybridized with p λ H15A (Fig. 4A), both RFLPs and a marked difference in the copy number of sequences hybridizing to p λ H15A were detected among DNAs derived from liver, lung, brain and other organs. Some caution is necessary in interpreting the observed pattern of RFLP because this patient had received combination chemotherapy of doxorubicin, vincristine, cytosine arabinoside and prednisone for 4 days. Nevertheless, we observed RFLPs in these tissues following digestion with *Eco*RI, *Bam*HI and *Hind*III which further indicated that tissue-related DNA rearrangements had occurred. A gene copy number reconstitution experiment showed that more than 50 copies of a DNA sequence hybridizing to p λ H15A are present in lung and brain tissues, whereas less than 10 copies are present in spleen and kidney tissue (data now shown). Finally, RFLP was also detected in the DNA derived from different organs of a 40-yr old female who had died of melanoma (Fig. 4B), again indicating a frequent occurrence of tissue-related sequence rearrangement using the p λ H15A probe.

Certain sequences hybridizing to p λ H15A occur as extrachromosomal molecules

Variations in the copy number of sequences hybridizing to p λ H15A may indicate that these sequences occur in extra-

chromosomal form and are differentially amplified in different tissues or physiological states. Initially, we observed a major band ~4.8 kb long in most of the DNAs examined after digestion with *EcoRI*, *BamHI*, *HindIII* and *BglII* (data now shown) and blot hybridization with λ H15A. Digestion with *PstI* revealed a single band of DNA that was ~150–200 bp smaller and probably represents cleavage in a single region where two or more *PstI* sites are closely spaced.

The very unusual finding of a single 4.8-kb fragment for each of the restriction enzymes again suggested that molecules containing sequences complementary to λ H15A may exist in extrachromosomal circular form. We examined this possibility by blot hybridization of undigested leukocyte DNA of seven normal donors with λ H15A, and detected bands in positions that did not correspond to the bulk of undigested chromosomal DNA. (Compare the ethidium bromide staining pattern with the blot hybridization pattern in Fig. 5A.) The predominant hybridizing bands migrate faster than the bulk of linear chromosomal DNA and correspond to closed circular (A) and nicked circular (B) forms of duplex DNA. A lower concentration of linear duplex DNA of this molecular weight is detected as a faint band migrating slightly faster than nicked circular DNA (labelled 4.8 kb in Fig. 5A, B). These three hybridizing bands occupy relative positions in the gel that would correspond to the positions of nicked circular, linear and covalently closed circular forms of 4.8-kb duplex DNA in similar conditions (ionic strength and voltage gradient) of agarose gel electrophoresis¹¹. Furthermore, when leukocyte DNA from a normal individual was subjected to a time course digestion with *BamHI* (Fig. 5B), both closed circular (A) and nicked circular forms were converted to the linear duplex form (labelled 4.8 kb). Note that the nicked circular and linear duplex forms of this DNA species are in low concentration in the undigested sample (Fig. 5B, first lane from the right).

Additional discrete bands of DNA hybridizing with λ H15A were detected (Fig. 5A). One of them seems to be at an intermediate position between the closed circular and nicked circular forms of the predominant bands of hybridizing DNA, while the remaining bands seem to migrate more slowly than the nicked circular form of the predominant species (bands above B). These additional bands may represent closed circular, linear and nicked circular duplex forms, respectively, of at least

two other species of extrachromosomal DNA that contain sequences complementary to λ H15A but that are present at much lower copy numbers. Perhaps one, but not both, of these species represents a circular oligomer, possibly a circular dimer form, of the predominant species (A and B in Fig. 5A). Furthermore, the time course of digestion of these DNAs with *BamHI* (Fig. 5B) converts all but one of the higher molecular weight DNA bands to 4.8-kb linear duplex DNA. This remaining band of DNA (C in Fig. 5B), therefore, must represent one additional closed circular DNA species containing sequences complementary to λ H15A.

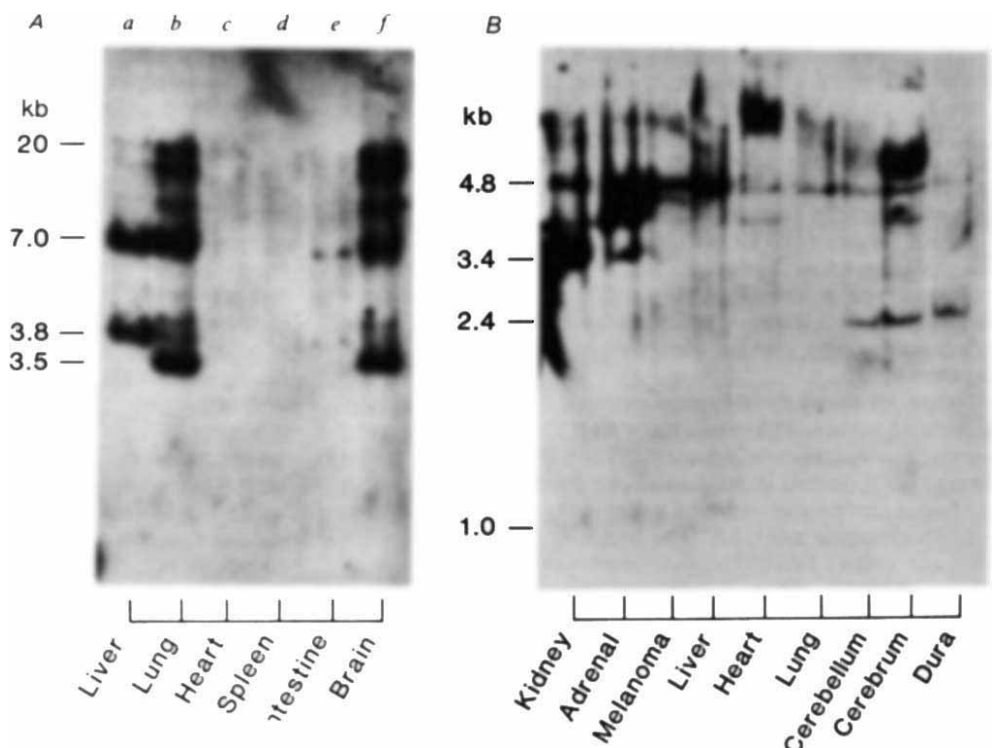
To visualize circular molecules containing sequences that hybridize to λ H15A, the undigested chromosomal DNA from leukocytes of a normal individual was fractionated by velocity sedimentation through a sucrose gradient. Fractions were collected and an aliquot of each fraction was blotted and hybridized with λ H15A DNA. One fraction enriched in nicked circular molecules hybridizing to λ H15A was extensively dialysed against TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0), concentrated to a final volume of 50 μ l and examined by electron microscopy using the aqueous Kleinschmidt technique¹².

As Fig. 5C shows, very low frequencies of circular duplex molecules were detected. Of 44 circular duplex molecules examined in this random sampling, 27% comprised a relatively uniform size class (upper row in Fig. 5C), with an average length of 4.71 ± 0.25 kb. A second, smaller homogeneous species was also detected (59% of all molecules), having an average length of 1.77 ± 0.35 kb. The remainder of circular duplex molecules appeared to be heterogeneous in size, with lengths ranging over 0.9–8.3 kb (lower row in Fig. 5C). The finding of a class of circular duplex molecules with a length of 4.7 kb is expected based on the results of blot hybridization experiments with λ H15, which had indicated the existence of a 4.8-kb circular duplex molecule. Supercoiled molecules (not shown) were detected in a faster sedimenting fraction of the sucrose gradient using the formamide modification¹² of the Kleinschmidt technique for electron microscopy.

Discussion

It is remarkable that *Alu* repeat sequences can exist in a clustered arrangement in the human genome, as revealed by the detailed analyses of the recombinant clone λ H15. The

Fig. 4 A, hybridization of ³²P-labelled λ H15A with *EcoRI*-digested DNA from different organs of an individual with acute lymphocytic leukaemia. At autopsy, several organs were obtained from a 22-yr old woman with ALL. DNA extraction restriction enzyme digestion, gel electrophoresis, Southern transfer and blot hybridization were performed as described in Fig. 2 legend. B, hybridization of ³²P-labelled λ H15A to *BamHI*-digested DNA derived from different organs of an individual with melanoma. At autopsy several organs were obtained from a 40-yr old woman with melanoma. The experiment was performed as described in Fig. 2 legend.



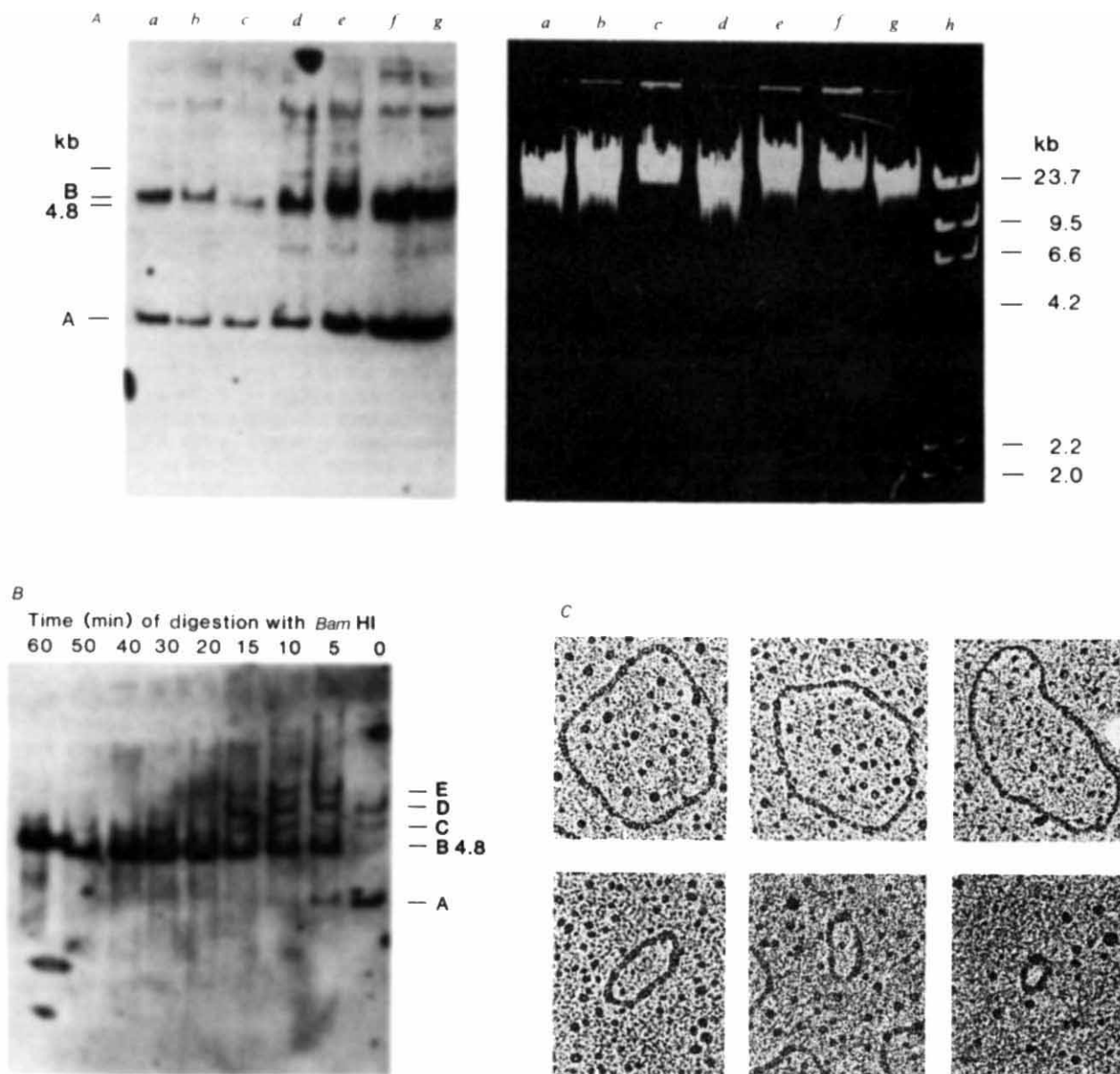


Fig. 5 A, hybridization of ³²P-labelled pAH15A to undigested DNA derived from leukocytes of normal donors. Undigested DNA (5 μg) from seven normal donors was electrophoresed through 1% agarose gels for ~15 h at 40 V. Southern transfer and blot hybridization were performed as described earlier. Hybridization of ³²P-labelled 0.8-kb human insert, purified by two cycles of gel electrophoresis, gave the same bands and relative intensities as blotting with the entire plasmid. B, time course of BamHI digestion of normal leukocyte DNA and hybridization with ³²P-labelled pAH15A. DNA (40 μg) was digested with BamHI at 1 U per μg and 5-μg aliquots were removed at different times. Gel electrophoresis, Southern transfer and blot hybridization with pAH15A were performed as described previously. C, electron micrographs of nicked circular duplex DNAs detected at very low frequency in aqueous Kleinschmidt spreads of nuclear DNA. Undigested DNA (100 μg) from leukocytes of a normal donor was applied to a 10–40% sucrose gradient buffered with 1 mM NaCl, 10 mM Tris-HCl pH 8.0 and 10 mM EDTA, and centrifuged in a VTi 50 rotor at 50,000 r.p.m. for 3 h. Thirty-six fractions of 1.1 ml each were collected and 100 μl of each fraction were electrophoresed in a 1% agarose gel, transferred to a nitrocellulose filter and hybridized with pAH15A as described earlier. The fraction enriched in molecules hybridizing to pAH15A was dialysed extensively against TE buffer, concentrated to a final volume of 50 μl and spread for electron microscopy using the aqueous Kleinschmidt technique¹². The molecules presented in the upper row have an apparent homogeneous length of 4.7 kb and represent ~27% of all circular duplexes detected. Other circular duplexes detected were smaller and heterogeneous in size, as shown in the lower row of micrographs, except for one species having an average length of 1.77 ± 0.35 kb. Length measurements were made relative to the bacteriophage PM2 DNA with a length of 10,260 bp. Scale bar, 1 μm.

clustered arrangement of *Alu* repeat sequences in λH15 may also be common to other regions of the human genome. We have screened the human genomic library with pAH15A, which contains a subcloned fragment of λH15, to detect DNA regions contiguous with the segment known as H15. Thus far, we have detected nine clones and have characterized five of them in detail. We observed that *Alu* repeat sequences are abundant in all five clones (T. Somasundaram *et al.*, unpublished) which reveals the existence of blocks of *Alu* sequences in the human genome. This fact is particularly intriguing with regard to the mechanism by which the pattern of interspersal observed for the bulk of *Alu* sequences is generated.

Sequence analysis showed that the *Alu* family sequence has

a dimeric structure with a duplication of a region conserved in mammalian evolution¹³. This region, 40 nucleotides long, is present once in the B1 family¹⁴ of repetitive sequences in mouse, which is analogous to the *Alu* family in man. This internal duplication could have arisen if the initial expansion of *Alu* sequences occurred in a clustered arrangement like that presented here. We may then visualize the pattern of interspersal of *Alu* repeat family members as arising from the sequential excision of *Alu* sequences and their subsequent insertion elsewhere in the genome by mechanisms involving nonhomologous or paralogous recombination as proposed by Anderson *et al.*¹⁵. The pattern of interspersal may have been fixed in evolution, with certain *Alu* repeat members having been

recruited for specific cellular functions, for example, in the initiation of DNA replication⁵ and as promoter sites for RNA polymerase III^{5,16}.

The discovery of a clustered arrangement of *Alu* repeat sequences provided an opportunity to explore the general question of sequence rearrangements in the human genome, especially those that may involve *Alu* repeat sequences. Using sequences from λ H15 DNA, we have detected two distinct aspects of DNA rearrangement involving either differential copy number or RFLPs. There were significant variations in copy number of sequences hybridizing to λ H15A among and between normal and leukaemic individuals. Furthermore, some of this variation could be tissue related, as the same set of sequences was present at different concentrations in different tissues derived from a single individual (Fig. 4A). Thus, sequences represented in the fragments hybridizing to λ H15A are not subject to selective pressures that would fix their number in the human genome.

It is important to emphasize that the 0.8-kb *Eco*RI fragment, which is the insert fragment in λ H15A, is itself embedded in a cluster of *Alu* repeat sequences in the human genome. The occurrence of differential copy number of this sequence could, therefore, be due to the increased frequency of recombination events involving repetitive sequences as proposed by Smith¹⁷ as well as amplification of extrachromosomal circular DNA forms. This latter phenomenon was clearly detected.

When total undigested DNA was electrophoresed and probed with λ H15A, we were surprised to find three hybridizing bands that resembled the gel electrophoretic separation of nicked circular, linear and closed-circular duplex forms of DNA (Fig. 5A). Circular duplex DNAs were also detected by electron microscopy after velocity sedimentation of total DNA, although there were two major species with lengths of ~ 4.7 and 1.8 kb respectively. Circular molecules with lengths other than 4.8 kb and hybridizing to λ H15A were detected in DNAs derived from sources other than leukocytes. A human skin fibroblast cell line after serial passage for eight generations was found to contain an additional species ~ 7.5 kb long. The DNAs from liver, lung and brain tissues of one individual (Fig. 4A) all contained extrachromosomal sequences present as circular molecules with sizes different from the predominant circular species detected in normal leukocytes. In particular, electron microscopic examination of the covalently closed circular DNA from lung tissue using propidium diiodide–CsCl density gradient centrifugation revealed a major population ($\sim 60\%$) of the circular DNAs having an average length of 3.4 ± 0.2 kb. In Southern blot analysis, λ H15A hybridized to this closed circular DNA population after *Eco*RI digestion and revealed one major band at 3.5 kb, essentially the same as identified by electron microscopy (B.C. *et al.*, unpublished). The DNA from kidney tissue of another individual contained additional hybridizing sequences corresponding to 4.0-kb circular duplex molecules (B.C. *et al.*, unpublished). These variations in size could reflect different genome rearrangements in formation of these circular molecules and may possibly involve duplications or deletions of *Alu* repeat sequences.

Blot hybridization of undigested leukocyte DNA with a cloned *Alu* repeat sequence, BLUR 8, or human repetitive DNA revealed the presence of *Alu* repeat sequences in the leukocyte 4.8-kb circular DNA (B.C. *et al.*, unpublished). In this case, we detected several discrete bands, one of which represented very low molecular weight DNA in the size range corresponding to the length of the *Alu* repeat sequence. This observation may indicate that *Alu* repeat sequences themselves occur as small circular duplex DNAs.

Small polydisperse circular DNAs have also been reported to occur at different concentrations in different organs of chicken embryos as well as in the chicken bursa at various stages of development¹⁸. It is possible that the difference in copy number of circular DNAs detected in the leukocyte DNA of normal and leukaemic individuals could reflect the different stages of differentiation of the leukocyte populations. Small

polydisperse circular DNAs were first described for HeLa cells but were found to be quite heterogeneous in size, with a mean length of 0.32 μ m (refs 19, 20). Interestingly, the small polydisperse circular DNAs in HeLa cells were shown to be of cytoplasmic but not mitochondrial origin²⁰. It is plausible that low concentrations of small polydisperse circular DNAs exist in the nucleus and, in the tissues we examined, particular size classes were selectively amplified. Note that the copy number of the 4.8-kb circular DNA we describe (~ 50 copies per leukocyte haploid genome) corresponds approximately to the copy number of the entire population (50–200 molecules per cell) of small polydisperse circular DNAs in HeLa cells²⁰, implying selective amplification of a particular size class in the population of polydisperse DNA examined. The circular DNAs we detected more closely resemble the small circular DNAs in nuclei of cultured cells of *Drosophila melanogaster*, which have been shown to contain middle-repetitive sequences²¹.

Although sequences hybridizing to λ H15A DNA occurred as extrachromosomal elements in the different tissues of a woman with leukaemia (Fig. 4A), we were unable to detect, unequivocally, extrachromosomal molecules without restriction enzyme digestion in tissues from a woman with melanoma (Fig. 4B). However, DNA from tissues of this individual contained a 4.8-kb hybridizing band (Fig. 4B) which in other cases corresponded to extrachromosomal DNA. The other discrete bands in the different tissues of this individual may indeed represent RFLP of chromosomal sequences. Therefore, the occurrence of H15A hybridizing restriction fragments of different lengths in different tissues of the same individual may involve either chromosomal or extrachromosomal sequences. Furthermore, the copy numbers and arrangement of sequences hybridizing to λ H15A as chromosomal segments appear to differ in other tissues. In this regard, a 3.4-kb *Bam*HI fragment (Fig. 3A, B) is present in three copies in leukocyte DNA whereas at least six copies are present in kidney tissue (Fig. 3B), indicating that a duplication involving this segment has occurred to different extents in the two cell types. The occurrence of such variation in copy number is further confirmed by the finding that nine clones contain sequences homologous to λ H15A in a human genomic DNA library established from fetal liver tissue⁷. Five of these clones were analysed and shown to represent segments of the human genome different from H15. Furthermore, in the liver tissue from which the library was established, the arrangement of sequences probed with λ H15A differs from that found in leukocyte DNA. This conclusion is drawn from the observation that possibly two chromosomal fragments in *Pst*I-digested leukocyte DNA hybridize with λ H15A (Fig. 2) rather than five or more as expected.

These combined studies indicate that polymorphisms and variation in copy number of chromosomal and extrachromosomal DNA sequences hybridizing with λ H15A reflect the presence of transposable elements in human cells. The recent observation that *copia*, the well known transposable element in *Drosophila*, exists in extrachromosomal form in cultured cells²² adds significance to the present results. In general, the human genome seems to be a dynamic structure in which variations can be introduced by sequence rearrangement, certain of which can lead to the formation of circular duplex DNA molecules. This genetic plasticity is quite characteristic of transposable elements, and the consequent genome alterations are relevant to evolutionary changes, while the DNA rearrangements may be involved in human cancer²³.

We thank Dr R. Shmookler Reis for many helpful suggestions during this work, Nancy Yen and Debbie Brock for technical assistance, T. Maniatis for the human genomic DNA library, P. Deininger and C. Schmid for the *Alu* repeat plasmids, and Abby Maizel for providing CEM cells. This investigation was supported by NIH grants GM 23965, GM 27774, CA 20124, CA 16672 and CA 16527, and Robert A. Welch Foundation grants G-267 and G-841. H.A.B.-S. was supported by a scholarship from the Consejo Nacional de Ciencia y Tecnología (CONACYT) of the Mexican government.

Received 25 August 1981; accepted 25 January 1982.

1. Rinehart, P. F., Ritch, G. T., Deininger, L. P. & Schmid, W. C. *Biochemistry* **20**, 3003–3010 (1980).
2. Tashima, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 1508–1512 (1981).
3. Houck, C. M., Rinehart, F. P. & Schmid, C. W. *J. molec. Biol.* **132**, 293–306 (1979).
4. Rubin, C. M., Houck, C. M., Deininger, P. L., Friedman, T. & Schmid, C. W. *Nature* **284**, 373–374 (1980).
5. Jelinek, W. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1390–1402 (1980).
6. Calabretta, B., Roberson, D. L., Maizel, A. L. & Saunders, G. F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6003–6007 (1981).
7. Lawn, R. W., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157–1174 (1978).
8. Kafatos, C. F., Jones, C. W. & Efstradiatis, A. *Nucleic Acids Res.* **7**, 1541–1552 (1979).
9. Shea-Griffin, R., Thireos, G., Kafatos, C. F., Petri, H. W. & Villa-Kamaroff, L. *Cell* **19**, 915–922 (1980).
10. Harper, M. E., Barrera-Saldana, H. A. & Saunders, G. F. *Am. J. hum. Genet.* (in the press).
11. Johnson, P. H. & Grossman, L. I. *Biochemistry* **16**, 4217–4225 (1977).
12. Davis, R. W., Simon, M. N. & Davidson, N. *Meth. Enzym.* **21**, 413–428 (1971).

13. Deininger, L. P., Jolly, J. D., Rubin, M. C., Friedmann, T. & Schmid, W. C. *J. molec. Biol.* **151**, 17–33 (1981).
14. Krayev, A. S. *et al. Nucleic Acids Res.* **8**, 425–440 (1980).
15. Anderson, D. M. *et al. J. molec. Biol.* **145**, 5–28 (1981).
16. Elder, J. T., Pan, J., Duncan, H. C. & Weissman, M. S. *Nucleic Acids Res.* **9**, 1171–1188 (1981).
17. Smith, G. P. *Science* **191**, 528–535 (1976).
18. De Lap, R. & Rush, G. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5855–5859 (1978).
19. Radloff, R., Bauer, W. & Vinograd, J. *Proc. natn. Acad. Sci. U.S.A.* **57**, 1514–1518 (1967).
20. Smith, C. A. & Vinograd, J. *J. molec. Biol.* **69**, 163–178 (1972).
21. Stanfield, S. W. & Lengyel, J. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6142–6146 (1979).
22. Flavell, J. A. & Ish-Horowitz, D. *Nature* **292**, 591–595 (1981).
23. Cairns, J. *Nature* **289**, 353–357 (1981).
24. De Wet, S. R. *et al. J. Virol.* **33**, 401–410 (1980).
25. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
26. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 32–38 (1973).
27. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1973).
28. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
29. Harper, M. & Saunders, G. F. *Chromosoma (Berl.)* **83**, 431–439 (1981).

LETTERS TO NATURE

A new model for white dwarf supernovae

R. Canal, J. Isern & J. Labay

Departamento de Física de la Tierra y del Cosmos,
Universidad de Barcelona, Barcelona 7, Spain

We show here that carbon–oxygen white dwarfs in close binary systems are likely progenitors of neutron stars and Type I supernovae (SNI). This conclusion is supported by different effects associated with crystallization of the carbon–oxygen mixture at relatively low temperatures ($T \leq 5 \times 10^7$ K) and high densities ($\rho \geq 5 \times 10^9$ g cm $^{-3}$). In particular, the existence of a pronounced eutectic in the phase diagram of carbon–oxygen mixtures induces carbon–oxygen separation, with oxygen settling at the star's centre and carbon being rejected to lower-density layers. This not only favours 'quiet' or non-explosive collapse, but it may also induce off-centre ignitions at variable depths. We discuss their possible correspondence with the 'fast' to 'slow' SNI range.

Carbon–oxygen white dwarfs in close binary systems are generally thought to be at the origin of nova outbursts¹ and their significance for SNI explosions has been suggested^{2,3}. Current models of SNI light curves^{4,5} involve explosive carbon–oxygen burning with ejection of variable amounts (from 0.5 to 1.0 $M_{\odot}$) of ^{56}Ni . This burning is triggered in various ways by mass accretion from the companion in the binary system. Several mechanisms have been proposed: central carbon deflagration for accretion rates in excess of $4 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$ (ref. 6); helium detonation at the base of the accreted layer, for $1.10^{-9} M_{\odot} \text{ yr}^{-1} \leq \dot{M} \leq 4 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$ (ref. 7); finally, central carbon deflagration or low-depth helium detonation again, for $\dot{M} \leq 1 \times 10^{-9} M_{\odot} \text{ yr}^{-1}$, the former for white dwarf masses in excess of 1.2 $M_{\odot}$, the latter for lower white dwarf masses⁸. All of these hypotheses suggest complete disruption of the white dwarf, with no remnant left and ejection of all of its material, with variable compositions. Also, these models have central temperatures of a few hundred million kelvins and ignite at a central density of about 3×10^9 g cm $^{-3}$.

Our model^{9–11} involves white dwarfs which have cooled for a long time ($t \geq 9 \times 10^8$ yr), so reaching low temperatures ($T \leq 5 \times 10^7$ K) and high central densities ($\rho_c \geq 5 \times 10^9$ g cm $^{-3}$); this implies a rather high mass: $M \geq 1 M_{\odot}$. Such a situation is likely when the white dwarf's companion star is a low-mass main-sequence star. Then, the detached phase of the system can be very long: $t \geq 10^9$ yr.

The main difference between this and models mentioned earlier is that the carbon–oxygen mixture becomes solid at such temperatures and densities. While carbon ignition in the fluid phase is still an open problem which awaits a two-dimensional

hydrodynamical treatment, solidification introduces important new features.

First, carbon is ignited by pycnonuclear reactions driven by the zero point energy. So, the thermal stability limit is increased to $\rho_c = 6 \times 10^9$ g cm $^{-3}$. This is still less than the dynamical stability limit, due to electron captures on ^{16}O , which is $\rho_c = 1.92 \times 10^{10}$ g cm $^{-3}$. But an important effect of solidification is to reduce the mechanisms of burning propagation (in the absence of detonation) to conduction alone¹². Because a detonation seems unlikely to form on carbon ignition for densities $\rho \geq 10^7$ g cm $^{-3}$ (ref. 13), solidification thus implies relatively low speeds for burning propagation ($\sim 10^{-5}$ times the speed of sound).

To exclude helium detonation in the outer, accreted layers, we limit ourselves to only two accretion regimes: fast accretion

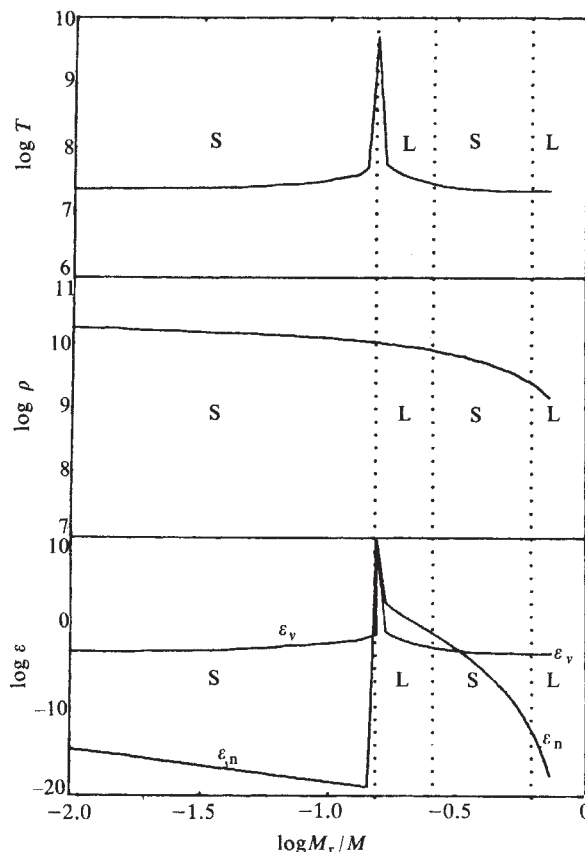


Fig. 1 Stellar structure during the off-centre ^{12}C flash. Temperature and density profiles are plotted against stellar mass fraction. Also shown are the energy releases (ϵ) by the nuclear (fusion) reactions (ϵ_n) and the e^- captures, together with the thermal neutrino losses (ϵ_v). S and L correspond to the solid and to the melted (Coulomb liquid) zones, respectively.

($\dot{M} \geq 10^{-8} M_{\odot} \text{ yr}^{-1}$), possibly associated with Roche lobe overflow by the companion, and slow accretion ($\dot{M} \leq 10^{-10} M_{\odot} \text{ yr}^{-1}$), more likely corresponding to accretion from the companion's stellar wind or to decay of the orbit by gravitational radiation.

We have also considered¹² a very fast accretion rate, $\dot{M} \sim 10^{-6} M_{\odot} \text{ yr}^{-1}$ (of the order of the Eddington limit for those stars), and obtained central ^{12}C ignition at $\rho_c = 1.37 \times 10^{10} \text{ g cm}^{-3}$. Due to the relative slowness of burning propagation in the solid phase, electron captures on the incinerated material overcome the effects of explosive burning and collapse ensues. This study did not include the behaviour of the accreted matter but only its effects in increasing the star's mass and density, and adiabatic gravitational compression. Problems such as the possible formation of an extended red giant envelope¹⁴ or heating of the core by a helium burning shell were not considered. So the real outcome remains unclear and this case serves only to illustrate one of the effects of solidification on pre-supernova evolution. For slower accretion rates those problems may be overcome but ^{12}C ignition happens at lower densities ($6 \times 10^9 \text{ g cm}^{-3} \leq \rho_c \leq 1 \times 10^{10} \text{ g cm}^{-3}$) and electron captures on the incinerated material are also slower. Those cases need further study.

A more important effect of solidification is the possibility of chemical separation. Stevenson¹⁵ has pointed out that the phase diagram of ^{12}C - ^{16}O mixtures presents a pronounced eutectic (or temperature minimum in the phase diagram) which corresponds to a composition of 66.8% ^{12}C and 33.2% ^{16}O , both by number. For a 'standard' chemical composition ($X_{\text{C}} = X_{\text{O}} = 0.50$), the solidification process would undergo the following: (1) ^{16}O 'snowflakes' settling at the centre of the star. (2) The oxygen-depleted fluid is rehomogenized by 'salt finger'-like instabilities. (3) Solid ^{12}C begins to form when the fluid composition becomes eutectic. (4) ^{12}C crystals, being lighter than the fluid mixture¹⁶, rise and redissolve higher up. (5) Given enough time, a completely differentiated body is formed.

Note that the cooling process is slowed down by the chemical separation process, as the last involves releasing gravitational potential energy. This effect can be estimated as follows.

The relative change in gravitational potential energy is of the order: $\Delta\Omega/\Omega \sim 10^{-3}$. Typical values for a $1 M_{\odot}$ star at the beginning of crystallization¹⁷ are: $-\Omega = 4.72 \times 10^{50} \text{ erg}$ and $L = 1.6 \times 10^{-3} L_{\odot}$. Thus $t = -\Delta\Omega/L = 2.5 \times 10^9 \text{ yr}$. This means a non-dramatic increase of the time scale of cooling, which was already $t_{\text{cool}} \sim 10^9 \text{ yr}$. It is appreciable, however, and this indicates that partial differentiation when mass accretion begins should be comparatively frequent.

We have calculated the precollapse and/or presupernova evolution of chemically differentiated carbon-oxygen white dwarfs. Our results show that for complete carbon-oxygen separation and for both slow and fast accretion rates (with the reserves previously pointed out concerning very fast rates), the behaviour of the central layers is: (1) heating of the star's centre by electron captures on ^{16}O , from $\rho_c = 1.92 \times 10^{10} \text{ g cm}^{-3}$. (2) ^{16}O melting and building up of a superadiabatic temperature gradient, with convection being stopped by the μ_c -gradient. (3) Growth of ^{16}C abundance through $^{16}\text{O} + 2e^- \rightarrow ^{16}\text{C} + 2\nu_e$. (4) ^{16}C ignition when its abundance reaches $\sim 10\%$ by number. The overpressures in the star's centre are small and cannot start a detonation. Burning propagates outwards by conduction alone (the surrounding layers are still solid). The electron captures on the incinerated material being very fast at those densities, ^{16}C ignition acts as the triggering mechanism for the collapse. Off-centre ignition of ^{12}C is still possible during the hydrodynamic compression of the outer half of the star's mass, but this phase is beyond the scope of the present study.

We have also evolved partially differentiated models, in particular those with a central oxygen core surrounded by an eutectic carbon-oxygen mixture ($\sim 0.25 M_{\odot}$ of ^{16}O surrounded by $\sim 1.15 M_{\odot}$ of ^{12}C - ^{16}O). For fast accretion the behaviour of these models is similar to that of the completely differentiated ones. For slow accretion ($\dot{M} = 10^{-12} M_{\odot} \text{ yr}^{-1}$), however, ^{12}C

ignition happens just at the base of the carbon-oxygen layers. The structure of this model at the time of carbon flash is shown in Fig. 1. Deflagrative carbon burning may, in this case, lead to the ejection of the carbon-oxygen layers after their incineration. The size of the central oxygen core being variable for different degrees of carbon-oxygen separation (corresponding to different cooling times), our model might account for thermonuclear burning of C-O, incineration and ejection of different amounts of material, in the approximative range of 0.5 – $1 M_{\odot}$. This would agree with Arnett's⁵ explanation of the 'fast' to 'slow' SN I range, which involves ejecting from $0.5 M_{\odot}$ of ^{56}Ni (in the 'slow' SN I) to $1 M_{\odot}$ of ^{56}Ni (in the 'fast' SN I).

Our model, besides leading to neutron star formation, seems also to be a promising explanation for Type I supernovae. The dynamical phase, of mass ejection and/or collapse of the inner zones, must be further explored to ascertain the exact amounts and the chemical composition of the matter expelled. Were our preliminary results confirmed, a single parameter (the duration of the cooling phase) would explain, within the slow accretion regime, the entire range of 'fast' to 'slow' SN I.

Received 23 November 1981; accepted 2 February 1982.

1. Starrfield, S., Truran, J. W., Sparks, W. M. & Kutter, G. S. *Astrophys. J.* **176**, 169 (1972).
2. Whelan, J. & Iben, I. *Astrophys. J.* **186**, 1007 (1973).
3. Canal, R. & Schatzman, E. *Astr. Astrophys.* **46**, 229 (1976).
4. Colgate, S. A., Petschek, A. G. & Kriese, J. T. *Astrophys. J. Lett.* **237**, L81 (1980).
5. Arnett, W. D. Preprint (Enrico Fermi Institute, University of Chicago, 1981).
6. Mazurek, T. J. & Wheeler, J. C. *Fundamentals cosmic Phys.* **5**, 193 (1980).
7. Woosley, S. E., Weaver, T. A. & Taam, R. E. in *Proc. Texas Workshop on Type I Supernovae*, 96 (1980).
8. Nomoto, K. in *Proc. Texas Workshop on Type I Supernovae*, 164 (1980).
9. Canal, R., Isern, J. & Labay, J. *Astrophys. J. Lett.* **241**, L33 (1980).
10. Canal, R., Isern, J. & Labay, J. *Space Sci. Rev.* **27**, 595 (1980).
11. Canal, R., Isern, J. & Labay, J. in *Proc. NATO Advanced Study Institute on Supernovae*, Cambridge (in the press).
12. Canal, R. & Isern, J. *Proc. IAU Colloq. No. 53*, 52 (1979).
13. Mazurek, T. J., Meier, D. C. & Wheeler, J. C. *Astrophys. J.* **213**, 518 (1977).
14. Paczynski, B. & Zytow, A. *Astrophys. J.* **222**, 604 (1978).
15. Stevenson, D. J. *J. Phys. Suppl.* **41**, C2-53 (1980).
16. Pollock, E. L. & Hansen, J. P. *Phys. Rev. A*, **8**, 3110 (1973).
17. Lamb, D. Q. & Van Horn, H. M. *Astrophys. J.* **200**, 306 (1975).

A Compton-cooled feedback mechanism for Cygnus X-1

P. W. Guilbert* & A. C. Fabian

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

The spectrum of the X-ray source Cygnus X-1 appears to exhibit two states. The source spends most of its time in the so-called 'low state' in which it is characterized by a power-law spectrum, $N(\epsilon) \propto \epsilon^{-\alpha}$, of photon index 1.6 and high-energy cutoff $\sim 100 \text{ keV}$ (ref. 1). Occasionally it makes a transition to a 'high state'² during which the soft flux (1–3 keV) increases by a factor of a few and the X-ray spectrum changes to an index steeper than 2.5. Most of the interpretation of the spectrum of Cyg X-1 has centred on the comptonization of soft photons by a hot gas in an accretion disk surrounding a black hole^{1,3}. It has been assumed that the hot gas is held at a fixed temperature, despite the obvious high degree of variability observed in the source on time scales ranging from milliseconds to months^{4,5}. We consider here the effect of relaxing this assumption of constant temperature because a Compton-cooled gas, in which the gas temperature drops in response to the comptonization of soft photons, can readily produce a two-state time-averaged spectrum from variations in the soft flux⁶. We speculate on a feedback mechanism that may effectively lock the source in either of the two states.

The slope and high-energy cutoff of the spectrum in the low state indicate, from computations of Compton-cooling gas⁶, that the Thomson depth τ of the gas is ≥ 2 and that the initial

* Present address: Joint Institute for Laboratory Astrophysics, University of Colorado and National Bureau of Standards, Boulder, Colorado 80309, USA.

gas temperature kT is ~ 100 keV. Page *et al.*⁷ measured a lag time (in the high state) for the variations in the 3–40 keV flux compared with the 1–3 keV flux of ~ 7.5 ms. This is characteristic of the shortest time scales seen in the source apart from the ill-defined submillisecond bursts, the significance of which may be in doubt⁵. Using this lag time scale as a rough guide to the photon escape time, t_e , from the hot gas we find

$$t_e \approx \frac{\tau^2}{n_e \sigma_{TC}} = \frac{\tau R_x}{c} = 7.5 \times 10^{-3} f \text{ s}$$

where n_e is the electron density, R_x is a typical dimension of a hot gas cloud in the source and f is a factor allowing for the uncertainty in cloud geometry. Thus

$$R_x = 1.1 \times 10^8 \left(\frac{2}{\tau}\right) f \text{ cm} \quad (1)$$

and

$$n_e = 2.7 \times 10^{16} \left(\frac{\tau}{2}\right)^2 f^{-1} \text{ cm}^{-3} \quad (2)$$

Conservation of energy then implies that the maximum X-ray luminosity

$$\begin{aligned} L_{\max} &\approx n_e (3kT) \frac{4\pi}{3} R_x^3 n_e t_e^{-1} \\ &= 10^{37} \left(\frac{2}{\tau}\right) n_e f \text{ erg s}^{-1} \end{aligned} \quad (3)$$

where n_e is the mean number of clouds present at a given time. The actual luminosity is $(t_e/t_{\text{cool}})L_{\max}$. (The maximum bremsstrahlung luminosity from this region is almost 100 times smaller.) Observations of L in the low state indicate that fn_e is at least a few^{1,2}. The average scattered photon energy is given by

$$\frac{\langle \epsilon_x \rangle}{\epsilon_0} = \frac{\alpha - 1}{2 - \alpha} \left[\left(\frac{kT}{\epsilon_0} \right)^{2-\alpha} - 1 \right] \left[1 - \left(\frac{\epsilon_0}{kT} \right)^{\alpha-1} \right]^{-1} \quad (4)$$

which in the low state, where $\alpha \sim 1.6$ and $\epsilon_0 \sim 0.5$ keV (refs 2, 8), gives $\langle \epsilon_x \rangle/\epsilon_0 = 12$. Then because the observed soft and hard luminosities are similar², we see that only $\sim 10\%$ of the soft photons can be scattered in the hot clouds. We have already noted that τ in the clouds > 2 , so we conclude that most of the soft photons originate in a region physically separate from the hot gas clouds, ruling out any simple coronal model. The cooling time of a cloud⁶

$$t_{\text{cool}} \approx \frac{3\epsilon_0}{L_s} \exp(t_{\text{therm}}/t_e) \frac{4\pi}{3} R_x^3 n_e, \quad (5)$$

where the time taken for photons to scatter to $\epsilon \sim kT$

$$t_{\text{therm}} = \frac{\log\left(\frac{kT}{\epsilon_0}\right)}{\log\left(1 + \frac{4K_3(x)}{xK_2(x)}\right)} (n_e \sigma_{TC})^{-1} \quad (6)$$

($x = m_e c^2/kT$, $K_3(x)$ and $K_2(x)$ are modified Bessel functions and L_s is the incident soft photon luminosity at ϵ_0). Now $L_s \sim 10^{36} n_e^{-1} \text{ erg s}^{-1}$ in the low state, so for $\tau = 2$ we find

$$t_{\text{cool}}^l \approx 2 \times 10^{-3} n_e f^2 \text{ s} \quad (7)$$

The hard spectrum (together with equation (3) which indicates the value of $n_e f$) implies that $t_{\text{cool}}^l > t_e$ in the low state. Large variations of the X-ray flux observed on time scales of 20–30 ms (ref. 5) indicate that $n_e f$ cannot much exceed 10, because a large average number of hot clouds would smooth out variations on time scales of a few times t_e . In the low state therefore, $t_e < t_{\text{cool}}^l < 3t_e$ as $\tau \geq 2$. The cooling time is inversely proportional to L_s , which increases in the high state, so that then $t_e > t_{\text{cool}}^h$. This means that the gas cools significantly before many photons are able to scatter far up in energy resulting in a steeper

spectrum⁶. Comptonized bremsstrahlung, from the rapidly cooling hot clouds, may produce a detectable hard tail to the spectrum in this state. Model spectra have been produced using the method given in ref. 9 assuming an initial gas temperature of 100 keV and predict greater X-ray flux above 100 keV than do the constant temperature comptonization spectra from gas at $kT \sim 30$ keV (refs 1, 10). We can obtain good agreement (by eye) over the 3–500 keV range of the HEAO 1 spectrum from a Compton-cooled model.

The time-averaged spectral index of Compton-cooled gas in the low state is independent⁶ of L_s provided that $t_{\text{cool}} > t_e$, although $L_x \propto L_s$. L_x can, moreover, change independently of L_s if τ , T or n_e change. The spectral index is sensitive to the soft flux in the high state, steepening (flattening) as the flux increases (decreases). The Compton-cooling model can be tested by observing the spectral evolution of a luminous burst. If the temperature of the hot gas is fixed, then the X-ray spectrum of the burst will continually harden with time until it assumes its final shape. If the gas cools, on the other hand, the spectrum will only harden until $t \sim t_{\text{cool}}$ (or t_e , which ever is the larger) and then it will begin to soften again. The minimum time scale for variations of the soft flux is expected to be the dynamical time scale (of the disk) in the region of its emission³

$$P \approx 0.5 \left(\frac{M_*}{M_\odot} \right)^{-1/2} \left(\frac{R}{10^8 \text{ cm}} \right)^{3/2} \text{ s}$$

We envisage a simplified picture of Cyg X-1 in which soft photons cool clouds of hot gas from $kT \sim 100$ keV on time scales of ~ 10 ms. The published spectra are the time averages from many clouds which may have passed through many heating (by shocks perhaps) and Compton-cooling phases. The system may lock itself in the two states through feedback provided by the X-ray spectrum. We suppose that the soft photon source is cool gas near to the hard X-ray emitting clouds. The soft photon flux determines t_{cool} and thus the slope of the spectrum of the X rays incident on the cool gas. When $\alpha > 2$ (the low state) the highest-energy photons (~ 100 keV) dominate this effect which is predominantly heating. This may reduce the soft flux and maintain $t_{\text{cool}} > t_e$ and thus $\alpha > 2$. When $\alpha < 2$ (the high state) the lowest-energy photons (> 1 keV) dominate and may even have a cooling effect on the soft photon-emitting gas. This may provoke the production of soft photons and maintain $t_{\text{cool}} > t_e$ and thus $\alpha > 2$. This situation therefore leads to the source latching into either state according to the initial value of L_s .

To be more specific, we envisage this process occurring in an accretion disk where the X-ray-emitting clouds are associated with a thickening due to the local heating exceeding the local cooling rate (the 'thermal instability'^{3,11}). The X rays add a term to either the local heating or cooling rate depending on the spectral slope, thereby increasing or decreasing the size of the X-ray-emitting region at the expense of the cool one. The importance of the X-radiation can be estimated by considering a boundary region at R_u of incremental radius h equal to the local half thickness of the cool disk. The flux of X-radiation incident on the boundary is $L_x h/R_u$, which is comparable with the local rate of dissipation of gravitational energy $GMMh/R_u^2$. This last quantity is $L_s^T h/R_u$ or $L_x h/R_u$, depending on whether the cool region is at a larger or smaller radius than the X-ray-emitting region. L_s^T is the total soft X-ray luminosity ($\sim L_x$ in the low state). An increase in $\dot{M}$, and thus of L_s , of sufficient magnitude to change t_{cool}/t_e , could change R_u and thereby force the system to change its state and remain there until an opposite change occurs. The inferred large value of L_s^T in Cyg X-1 ($L_s^T > L_x$) suggests, for an accretion model, that the luminous soft flux originates from an area at a smaller radius than the X-ray region (which requires some special geometry in order that only 10% of the photons are intercepted by the hot gas—see Fig. 2 in ref. 12 which illustrates this general configuration; the details are different). This would mean that soft (< 2 keV) X-ray measurements of Cyg X-1 are the most direct probes of

the black hole, and that the soft X rays probably control the shortest time-scale variability. Possible alternative solutions are that $\dot{M}$ decreases with radius due to a wind, or that the efficiency is low in the hot region because the black hole swallows much of the kinetic energy with the matter. Clearly our model requires a significant perturbation to drive it from one state to the other. The size of the necessary perturbation required to drive the system from low to high (high to low) will depend on $t_{\text{cool}}^{\text{h}}/t_{\text{e}}(t_{\text{cool}}^{\text{l}}/t_{\text{e}})$. The longer time spent in the low state may simply reflect the greater stability of that state to changes in the soft flux.

We conclude that, in the framework of a Compton-cooled plasma, the flat hard X-ray spectrum and rapid variability observed in the low state of Cyg X-1 imply that the Compton-cooling time, t_{cool} , just exceeds the photon escape time from the hot clouds, t_{e} . Feedback through the following loop then maintains that low state: $\alpha < 2 \Rightarrow$ emission from hot gas heats surrounding cool gas $\Rightarrow L_{\text{s}}$ decreases $\Rightarrow t_{\text{cool}}$ increases $\Rightarrow t_{\text{cool}} > t_{\text{e}} \Rightarrow \alpha < 2$. A substantial increase in L_{s} , perhaps due to a temporary change in $\dot{M}$, may then cause a change to the high-state loop: $\alpha > 2 \Rightarrow$ emission from hot gas cools cool gas $\Rightarrow L_{\text{s}}$ increases $\Rightarrow t_{\text{cool}} < t_{\text{e}} \Rightarrow \alpha > 2$.

A.C.F. thanks the Radcliffe Trust for support and P.W.G. the Guernsey States Education Council.

Received 29 September 1981; accepted 27 January 1982.

1. Sunyaev, R. A. & Trumper, J. *Nature* **279**, 506 (1979).
2. Chiapetti, L. *et al. Mon. Not. R. astr. Soc.* **197**, 139 (1981).
3. Eardley, D. M., Lightman, A. P., Shakura, N. I., Shapiro, S. L. & Sunyaev, R. A. *Comments Astrophys.* **7**, 161 (1978).
4. Nolan, P. L., Gruber, D. E., Matteson, J. L., Peterson, L. E. & Rothschild, R. E. *Astrophys. J.* **246**, 494 (1981).
5. Giles, A. B. *Mon. Not. R. astr. Soc.* **195**, 721 (1981).
6. Guilbert, P. W., Ross, R. R. & Fabian, A. C. *Mon. Not. R. astr. Soc.* (in the press).
7. Page, C. G., Bennetts, A. J. & Ricketts, M. J. *Space Sci. Rev.* **30**, 369 (1982).
8. Bedford, D. K. *et al. Space Sci. Rev.* **30**, 373 (1982).
9. Guilbert, P. W. *Mon. Not. R. astr. Soc.* **197**, 451 (1981).
10. Nolan, P. L. *et al. Nature* **193**, 275 (1981).
11. Pringle, J. E. *A. Rev. Astr. Astrophys.* **19**, 137 (1981).
12. Pringle, J. E., Rees, M. J. & Pacholczyk, A. G. *Astr. Astrophys.* **29**, 179 (1973).

Alkylation of lead (II) salts to tetraalkyllead in aqueous solution

L. J. Snyder

Ethyl Corporation, R & D Laboratory, Baton Rouge, Louisiana 70801, USA

J. M. Bentz

Ethyl Corporation, Research Laboratory, Detroit, Michigan, USA

It was recently reported by Ahmad *et al.*¹ that lead (II) salts react with alkyl iodides in aqueous solution to produce tetraalkyllead compounds. They stated this reaction would have serious environmental consequences for water quality and for the well-being of aquatic biota. The experiments were carried out in a small (250 ml) flask nearly filled with an aqueous lead (II) acetate solution (210 ml) and sealed with an aluminium foil wrapped stopper. We have now repeated the experiment and report that tetramethyllead (TML) is not formed by direct reaction of lead (II) salts with methyl iodide in an aqueous system, but comes from a secondary reaction of metallic lead with methyl iodide^{2,3}. Apparently, Ahmad *et al.*¹ inadvertently contacted the aluminium foil-wrapped stopper with a small quantity of lead (II) acetate solution to produce finely divided metallic lead which reacted with methyl iodide to produce TML. The reduction of lead (II) salts to metallic lead in both acid and alkaline solutions by aluminium and other metals has been reported previously⁴. When the reactions described by Ahmad *et al.*¹ were carried out in the absence and presence of alu-

minium, TML formed in those tests in which lead (II) acetate solution contacted aluminium but not when aluminium was absent. Our results indicate that methylation of lead (II) acetate does not occur in the environment to form tetraalkyllead compounds which would be deleterious to water quality.

We have repeated the alkylation experiments described by Ahmad *et al.*¹ by reacting lead (II) acetate with methyl iodide in aqueous solutions adjusted to pH 5–6 and pH 13 respectively. The experiments carried out in alkaline solution at pH 13 are reported in Table 1. Alkylation runs 1 and 2 were carried out as described by Ahmad *et al.*¹ by dissolving 1.0 mmol of lead (II) acetate in 200 ml of deionized water in a 250 ml filtration flask fitted with an aluminium foil-wrapped stopper. Methyl iodide (0.1 ml) was added and the flask was sealed and stored in a closed cabinet. After 5 days the mixture (vapour space and aqueous solution) was extracted with *n*-heptane. The TML content of *n*-heptane was measured by a colorimetric dithizone method⁵ modified to measure the absorbance of dialkyllead dithizonate in a small volume (0.9 ml) of chloroformdithizone solution at 485 nm wavelength. Results from the first two experiments (Table 1) using an independent method of analysis agreed well with those reported by Ahmad *et al.*¹. In the next three runs, contact between the aluminium foil-wrapped stopper and the lead (II) acetate solution was controlled. The foil in runs 3 and 4 was deliberately and completely wetted with lead (II) acetate solution (pH 13) at the start and occasionally during the experiment. Special precautions were taken to avoid contact between the foil-wrapped stopper and lead (II) acetate solution in run 5. The results of these runs demonstrate that TML is produced only if the aluminium foil-wrapped stopper contracts the lead (II) acetate solution. Significantly more TML was formed when the aluminium foil was intentionally wetted than in runs 1 and 2 where there was minimal and unintentional contact between lead (II) acetate solution and aluminium foil.

As further proof that there is no reaction between lead (II) acetate and methyl iodide in aqueous solution (pH 13) to produce TML, we increased the sensitivity of the colorimetric dithizone method by a factor of 10 by scaling up proportionally the quantities of lead (II) acetate, methyl iodide, aqueous solution and glassware used in the alkylation reaction. The reactants and solvent were maintained in the ratios used by Ahmad *et al.*¹. Repeat experiments were carried out in the absence of aluminium with the bottles sealed with screw caps lined with Teflon gaskets. No detectable quantities of TML (<0.01 p.p.b. as Pb) were produced from these tests (runs 6–10) carried out over 5–50 days.

Methylation tests were then carried out at pH 5–6. Aluminium was added directly to the lead (II) acetate solution in

Table 1 Reaction of lead (II) acetate with methyl iodide in aqueous solution (pH 13)

Run no.	Reaction period (days)	Experimental conditions	TML found (p.p.b. Pb in water)
1	5	a	1.2
2	6	a	1.5
3	5	b	8.5
4	5	b	10.2
5	5	c	<0.1
6	5	d	<0.01
7	5	d	<0.01
8	10	d	<
9	18	d	<0.01
10	50	d	<0.01

a, Aluminium foil present and methylation test carried out as described by Ahmad *et al.*¹. b, Aluminium foil deliberately wetted with lead (II) acetate solution (pH 13). c, Aluminium foil present but kept free of lead (II) acetate solution (pH 13). d, Absence of aluminium foil; the methylation experiments were scaled up proportionally to provide increased analytical sensitivity. In all cases, temperature was $23 \pm 2^\circ\text{C}$. Note that Ahmad *et al.*¹ found 370.5 ng of TML in a flask containing 200 ml of aqueous solution, which is equivalent to a 1.44 p.p.b. calculated as Pb in water.

Table 2 Reaction of metallic lead with methyl iodide in aqueous solution

Run no.	Reaction period (days)	pH	TML found (p.p.b. Pb in water)
1	7	13	744.0
2	6	13	2,375.0
3	5	Initially 5.5	<0.01
4	9	Initially 5.5	<0.01

Sample no. 2 contained 10 g of freshly sliced metallic lead while sample 1 contained 5 g of lead. However, the weight of lead is less important than the surface area of the freshly sliced lead. Samples 3 and 4 contained 8 g of freshly sliced lead. All tests contained 1,140 p.p.b. of methyl iodide in water. Temperature $23 \pm 2^\circ\text{C}$.

some experiments and omitted in others. The head space in the 250 ml flask was analysed by a gas chromatography/atomic absorption spectrometry procedure essentially as described in ref. 6 and having the same sensitivity as the method of Ahmad *et al.*¹. In these experiments, TML was detected only when the aluminium was contacted with lead (II) acetate solution.

That metallic lead is produced from the reduction of lead (II) acetate was confirmed by adding aluminium foil to (1) lead (II) acetate dissolved in distilled water (pH 5.5) and (2) lead (II) acetate solution adjusted to pH 13 with NaOH. The metallic lead produced was identified by scanning electron microscopy/energy-dispersive X-ray spectrometry. The reduction at pH 5.5 is much slower than at pH 13, which may explain why Ahmad *et al.*¹ found significantly smaller quantities of TML in slightly acid solutions.

Although the reaction of metallic lead with methyl iodide has been reported previously^{2,3}, we have demonstrated here that this reaction depends on the extent and character of the lead surface and the specific conditions in which the test is carried out. When we added freshly sliced metallic lead to a NaOH solution (pH 13) containing methyl iodide, TML was readily produced (Table 2). On the other hand, no TML was detected when freshly sliced lead was added to distilled water (pH 5.5) containing methyl iodide and the mixture allowed to stand in the dark and in a completely sedentary condition for 5–9 days (Table 2). In the latter case a rapid growth of insoluble hydrated lead oxides formed on the surface of the metallic lead, presumably inhibiting the alkylation reactions. In the presence of surface waters, metallic lead would be expected to react with other anions, such as sulphides, carbonates and phosphates, to form water-insoluble lead compounds adhering to the surface of the metal that would reduce or eliminate any alkylation of metallic lead with alkyl iodides, even if they were present in very high concentrations. Caution must be exercised against predicting whether metallic lead may be methylated in the environment by extrapolating information from laboratory experiments, especially as metallic lead and methyl iodide are chemically unstable in water.

When Jarvie and Whitmore⁷ repeated the lead (II) salt–methyl iodide reactions described by Ahmad *et al.*¹, no TML was observed when the reactions were carried out in the presence and in the absence of aluminium foil-wrapped stoppers. These authors attempted to explain their negative results on the basis of possible minor differences in temperature, time, pH, light effects, for example, from the conditions used by Ahmad *et al.*¹. However, we believe the correct explanation is that none of the aqueous lead acetate solution inadvertently contacted the aluminium foil-wrapped stoppers.

Received 6 August; accepted 18 December 1981.

- Ahmad, I., Chau, Y. K., Wong, P. T. S., Carty, A. J. & Taylor, L. *Nature* **287**, 716–717 (1980).
- Shapiro, H. & Frey, F. W. in *Organic Compounds of Lead* (Wiley, New York, 1968).
- Cahours, A. C. *r. heb. Séanc. Acad. Sci., Paris* **36**, 1001 (1853); *Ann. Chem.* **62**, 257 (1861).
- Mellor, J. W. in *Comprehensive Treatise on Inorganic and Theoretical Chemistry* Vol. 7, 506.
- Snyder, L. J. US Patent no. 3,912, 454 (1975).
- Chau, Y. K., Wong, P. T. S., Goulden, P. D. *Analyt. Chim. Acta* **85**, 421–424 (1976).
- Jarvie, A. W. P. & Whitmore, A. P. *Envir. Technol. Lett.* **2**, 197–204 (1981).

Aerosols of high chlorine concentration transported into central and eastern United States

R. W. Shaw Jr & F. S. Binkowski

Environmental Sciences Research Laboratory,
US Environmental Protection Agency, Research Triangle Park,
North Carolina 27711, USA

W. J. Courtney

Northrop Services, Inc., Research Triangle Park,
North Carolina 27709, USA

During 18–21 December 1980 a remarkable change occurred in the composition of the aerosol over a broad region of the central and eastern United States: the weight per cent of chlorine present in small aerosol particles increased by nearly two orders of magnitude. We report here an analysis of the aerosol composition and meteorology during the period of change which suggests that this increase was the result not of a pollution episode, but of a massive intrusion of very clean maritime air from the northern Pacific into a region extending at least from Wisconsin southwards to Kansas and Missouri and eastwards to Ohio.

We obtained elemental composition data for the event from X-ray fluorescence analysis of aerosols collected using virtual impactor (dichotomous) samplers¹. These samplers separate aerosols into 'fine' and 'coarse' fractions (<2.5 and 2.5–15 μm in aerodynamic diameter, respectively). Particles in the fine aerosol fraction have settling velocities in still air of $\leq 10^{-2} \text{ cm s}^{-1}$; consequently they may remain suspended for days or longer and may be transported long distances by large-scale atmospheric motion. The fractions were collected on Teflon membrane filters over 24-h intervals starting at midnight. Locations where aerosol samplers were operating during the period of interest and for which we have aerosol elemental composition data are three rural sites (in Kentucky, Indiana and Ohio); Portage, Wisconsin; Topeka, Kansas; and St Louis, Missouri¹. The rural sites were chosen as being representative of regional air quality and to avoid local emission sources as far as possible; these sites were located at $37^\circ 41' \text{ N}$, $88^\circ 01' \text{ W}$; $39^\circ 28' \text{ N}$, $84^\circ 56' \text{ W}$; and $40^\circ 36' \text{ N}$, $82^\circ 19' \text{ W}$. Aerosols were collected on odd-numbered days in Portage and Topeka and daily elsewhere. We estimate the analytical precision for Cl^- aerosol concentrations based on the standard deviation of many blank samples to be 2 ng m^{-3} .

The chlorine concentrations are shown as a percentage of fine aerosol mass in Fig. 1. The ratios of chlorine mass on the peak day to average chlorine mass for several days before and after the event are: Portage (76), Topeka (48), St Louis (28), rural Kentucky (43), rural Indiana (45) and rural Ohio (15). The peak chlorine concentrations in the fine aerosol fraction exceeded $1 \mu\text{g m}^{-3}$ at all locations except the easternmost site in Ohio. These are the highest fine aerosol chlorine concentrations that we have found among thousands of samples collected in urban and rural areas across the United States during the past 4 yr.

Two previously identified sources of elevated aerosol chlorine are road salting and vegetative burning^{2,3}, although these seem unlikely explanations for an event covering such a large area, they can be ruled out for other reasons, including aerosol properties. Elevated chlorine levels have been observed in aerosols after salting of roads to remove snow. After road salting, most of the chlorine is found in the coarse fraction; during the event reported here, however, the fine to coarse aerosol chlorine ratio was $>2:1$. City records show that roads

were salted in Portage on 2, 5, 23 and 29 December; in Topeka during 8–11 December and in St Louis on 23 and 24 December. These dates do not correspond to the observed event.

Vegetative burning causes marked increases in fine aerosol potassium. During the chlorine event, however, the potassium levels did not increase significantly in rural areas and decreased in the most polluted area, St Louis. Burning causes large increases in aerosol carbon; soot is an effective light absorber and renders the aerosol black. During the chlorine event, however, the blackness of the collected aerosol, as measured by light transmission of samples from the rural Indiana site, decreased by nearly an order of magnitude. Finally, when the chlorine aerosol levels increased, the total fine aerosol mass decreased by a factor of 2–3 in those areas where the fine aerosol is usually high due to industrial, residential and vehicle emissions (that is the sites at St Louis and rural Kentucky, Indiana and Ohio); this indicates ventilation by clean air. Ion-exchange chromatography showed that the concentration of sodium in the samples from the rural Kentucky site began increasing on 19 December, reached a peak value on 20 December which was 10 times higher than the 17 and 18 December levels, and diminished on 21 and 22 December. The mass ratio of Cl^-/Na^+ on 20 December was 1.5, equal to the stoichiometric mass ratio for NaCl. These aerosol compositions are consistent with a large, very clean air mass having moved suddenly into the eastern United States, carrying fine aerosol sea salt.

The meteorological conditions associated with this chlorine event are best described as a typical wintertime cold outbreak. A large, cold, dry anticyclone moved over the area of interest and, by 07.00 h on 20 December, covered the United States east of the Rocky Mountains (all times here are Eastern Stan-

dard Time, EST). The large-scale sinking motion (subsidence) accompanying the cold outbreak brought mid- and upper tropospheric air to lower levels as indicated by an isentropic trajectory analysis⁴. Such a trajectory, plotted backwards in time from 07.00 h on 19 December from a point over Salem, Illinois, showed that air which was at 850 mbar over Salem, was at 548 mbar over Whitehorse in the southern Yukon 48 h earlier. A second, simultaneous trajectory from Green Bay, Wisconsin did not show so much sinking but remained near 700 mbar and 48 h earlier was over Hudson Bay. Note that the trajectory at Salem is for air arriving only a few hours after the frontal passage, while that at Green Bay is for air arriving more than a day after the frontal passage. The stronger subsidence occurring at Salem is consistent with the view⁵ that the strongest subsidence in a cold outbreak occurs shortly after the frontal passage. Thus, the air arriving over our aerosol samplers at 07.00 h on 19 December came from mid-tropospheric altitudes and was over western and central Canada 48 h before. This air was mixed down to the surface by turbulence in the planetary boundary layer. The surface winds in the area of interest were between 5 and 10 m s^{-1} on the day in question.

The chlorine concentrations reported here are less than those which have been observed in the marine boundary layer, but much larger than those observed in either the continental boundary layer or free troposphere^{6,7}. Consequently we infer that the aerosols are of marine origin. We have not ascertained the exact source of this air, but we speculate that the air was in contact with the open sea long enough to acquire marine characteristics, and was then transported by synoptic-scale motion to the observing sites. There is evidence that air does indeed move from surface to mid-tropospheric levels in frontal cyclones⁵, and we suggest that this rising motion, followed by motion described by the calculated trajectories, was responsible for our observations. Transport of clean air having maritime characteristics to a continental site is the reverse of the transport of continental crustal aerosol to remote marine sites⁸.

Although these observations of high, fine-particle chlorine concentrations are unique as far as we know, nothing in the meteorological conditions seems to distinguish this cold outbreak from others typical of the season. The explanation may be associated with the interesting behaviour of the sulphate component of the fine aerosols which we observed at the Indiana sampling site. Studies in urban areas and in rural areas of the eastern United States have shown that sulphate is generally the most abundant component in the fine aerosol and is largely in the form of ammonium acid sulphate or ammonium sulphate^{9–11}. Included at the Indiana site was an aerosol sulphate monitor which thermally separated and measured the relatively volatile and non-volatile sulphates in the fine aerosol at intervals of 30 min¹². The non-volatile sulphates are thought to be alkali metal and alkaline earth sulphates, whereas the volatile sulphates consist of ammonium sulphate and ammonium acid sulphate. Preceding the event at the Indiana site, the volatile to non-volatile sulphate ratio was ~2:1. Between 23.00 and 24.00 h on 18 December, the amount of volatile sulphate fell abruptly and the non-volatile sulphate decreased slightly, resulting in a volatile to non-volatile sulphate ratio of 1:4. On the night of 21 December, at the end of the chlorine event, the volatile sulphate began to rise to pre-event levels. It may be that, in most cases, chlorine is lost from marine aerosols transported over long distances inland by reactions with acid gases (nitric acid) and aerosols (sulphuric acid and ammonium acid sulphate). The reaction would release chlorine from the marine aerosol in the form of the volatile gas hydrogen chloride. In our observations the abrupt decrease in volatile sulphate before the chlorine event suggests that there was no reactive loss of chlorine from the incoming marine aerosols.

Thus we have observed a brief event of very high concentrations of fine-mode aerosol chlorine at six locations in the central and eastern United States. Presumably, this was the result of transport from the northern Pacific of marine aerosols in an air parcel which displaced aerosols characteristic of the region.

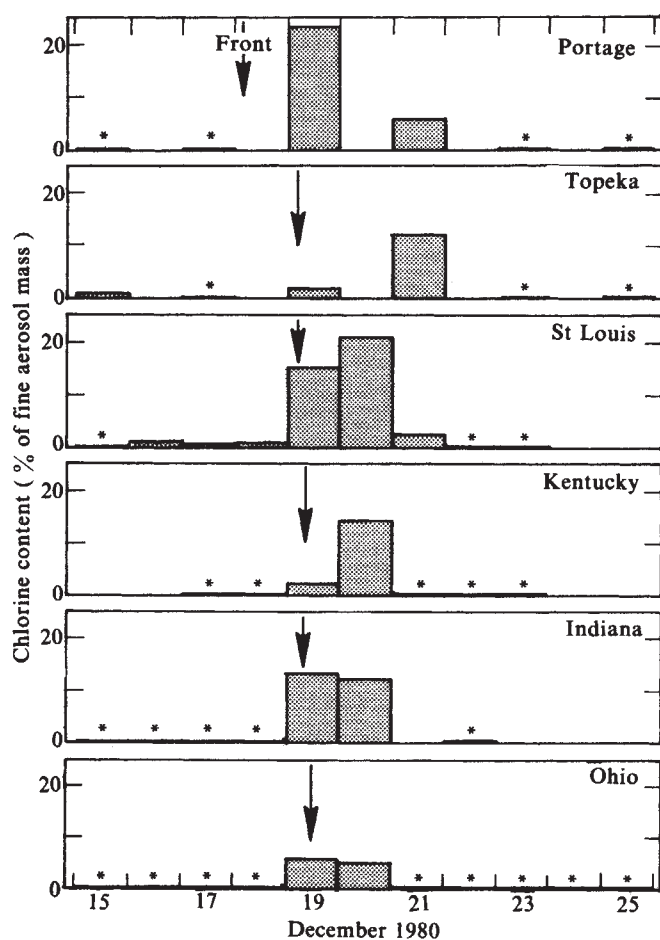


Fig. 1 Chlorine as percentage of fine-fraction aerosol mass collected during 24-h intervals. Note that samples were collected on odd-numbered days at Topeka and Portage. Asterisk indicates <0.5%.

We thank Philip Haagenson of the National Center for Atmospheric Research for the trajectory calculations, John Spengler and his colleagues at Harvard for supplying some of the aerosol samples, and James Huntzicker and colleagues of Oregon Graduate Center for providing sulphate data. Samples in rural Kentucky, Indiana and Ohio were collected by Mead Technology for the US Environmental Protection Agency. Samples from Portage, Topeka, St Louis were collected by the School of Public Health, Harvard University. F.S.B. was on assignment from the National Oceanic and Atmospheric Administration.

Received 25 November 1981; accepted 20 January 1982.

1. Dzuby, T. G. & Stevens, R. K. *Envir. Sci. Technol.* **9**, 663–668 (1975).
2. Dzuby, T. G., Stevens, R. K., Courtney, W. J. & Drane, E. A. *Electron Microscopy and X-Ray Applications* (ed. Russell, P. S.) **23** (Ann Arbor Science, Ann Arbor, 1981).
3. Watson, J. G. thesis, Oregon Graduate Center (1979).
4. Haagenson, P. L. & Shapiro, M. A. *Isentropic Trajectories for Derivation of Objectively Analyzed Meteorological Parameters* NCAR/TN-149+5TR (National Center for Atmospheric Research, Colorado, 1979).
5. Carlson, T. N. *Mon. Weath. Rev.* **108**, 1498–1509 (1980).
6. Huebert, B. J. & Lazrus, A. L. *J. geophys. Res.* **85**, 7337–7344 (1980).
7. Patterson, E. M. *et al. J. geophys. Res.* **85**, 7361–7376 (1980).
8. Duce, R. A., Unni, C. K., Ray, B. J., Prospero, J. M. & Merrill, J. T. *Science* **209**, 1522–1524 (1980).
9. Brosset, C. *Atmos. Envir.* **12**, 25–38 (1978).
10. Charlson, R. J., Covert, D. S., Larson, T. V. & Waggoner, A. P. *Atmos. Envir.* **12**, 39–53 (1978).
11. Stevens, R. K. *et al. Envir. Sci. Technol.* **14**, 1491–1498 (1980).
12. Huntzicker, J. J., Hoffmann, R. S. & Ling, G. S. *Atmos. Envir.* **12**, 83–88 (1978).

Thermoluminescence reveals weathering stages in basaltic rocks

Y. Göksu Ögelman* & S. Kapur†

* Department of Physics and Space Sciences, † Department of Soil Sciences, Çukurova University, Adana, Turkey

The weathering of basaltic rocks is important in several regions throughout the world because of its soil forming capacity^{1,2}. Basaltic minerals alter to various types of clay minerals depending on the climatic and chemical conditions of the environment. Investigation of the weathering of rocks involves detailed work consisting of thin section and X-ray and electron microscopy techniques: alternative methods would be desirable. We show here that the thermoluminescence (TL) properties of weathered feldspars can be used as tracers of the relative weathering stages of basaltic rocks.

TL techniques have been used successfully to date pottery^{3–5}, burned flints⁶, lava flows⁷, young ocean sediments⁸ and biological materials like bone and teeth⁹. Basically the technique involves the emission of visible light from heated crystalline material, as the electrons in traps are released and recombine with the holes. In minerals and rocks, the ionizing radiation, primarily from the naturally occurring radioisotopes of ²³⁸U, ²³²Th and ⁴⁰K, pumps the electrons into the traps. The TL efficiency and the radiation dose at which light emission of the material will saturate are a complex function of the ion vacancies, interstitial ions, impurities and dislocations that lead to the formation of traps with semicontinuous energy levels.

Using TL for measurement of geological age, however, has led to difficulties because of the complexity of the geological history involving structural stresses, recrystallization and non-thermal fading of high temperature peaks known as the anomalous fading¹⁰.

We initially tried to date samples collected from the parent materials of four sites in south-east Turkey, between Urfa and Siverek. However, their markedly different TL characteristics suggested that we should rather investigate their weathering properties. This region is composed of basaltic rocks of Tertiary age covering the northern part of Syria—the Syrian Shield. The

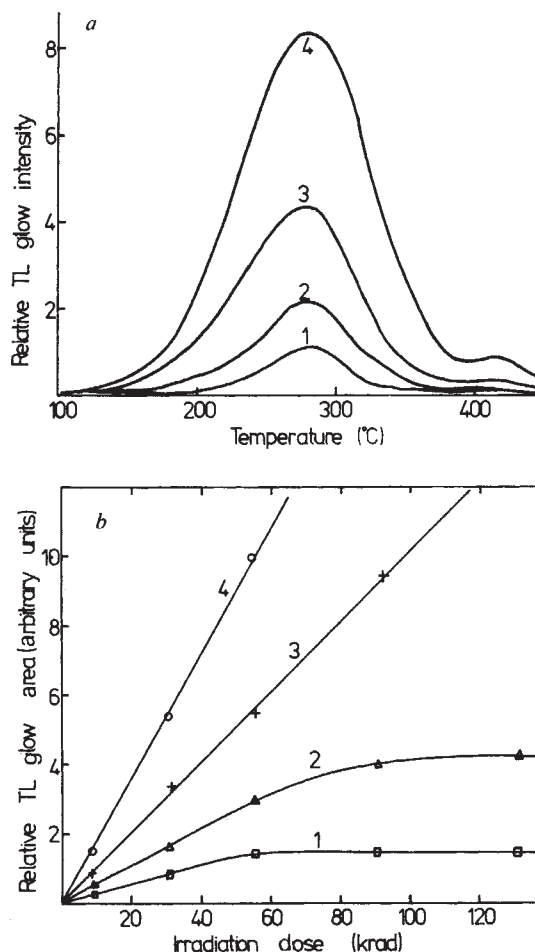


Fig. 1 *a*, Induced TL glow curves of weathered feldspar from four different profiles of the Karaca Mountain basaltic plateau. The irradiation dose was 55 krad and the heating rate 2°C s^{-1} . Low-temperature peaks were erased by preheating the samples at 80°C for 5 s. *b*, Variation of TL glow curve area under the 275°C peak as a function of the irradiation dose. The saturation doses were 55, 80, 130 and 300 krad respectively from layers 1 to 4. Points corresponding to high radiation doses of 3 and 4 are not shown. The numbers on the curves indicate their relative weathering in increasing order.

samples came from profiles typical of topographically and chronologically sequential continuations of different lava flow levels of the Karaca Mountain (1,919 m).

Feldspar grains, which are responsible for producing TL in basaltic rocks, were separated out from the samples. The basalts were viced in a V-shaped metal to extract feldspars without damaging them. They were then washed with HCl to dissolve the secondary CaCO_3 depositions in the vesicles. After dispersing the samples in Calgon overnight, grains of $100\text{--}150\text{ }\mu\text{m}$ size were separated by wet sieving. Lighter feldspar grains were extracted by floating them in bromoform. Aliquots of 5-mg feldspar grains were spread over 1-cm diameter aluminium disks for TL measurements. These samples were heated in an argon atmosphere from 100 to 480°C at a rate of 2°C s^{-1} and natural TL glow curves were obtained. It was observed that during this process, the natural TL of all the samples was almost completely drained. Subsequently, varying radiation doses were given to the samples using a $200\text{ mCi } ^{90}\text{Sr}/^{90}\text{Yr}$ β source and glow curves were obtained 5 min later. Figure 1*a* shows the induced TL glow curves of the feldspar samples from the four different profiles of the Karaca Mountain basaltic plateau. Figure 1*b* shows the variation of the TL glow area under the 275°C peak with radiation dose. Each point represents the average of four measurements from each sample. The TL sensitivities of the samples were obtained by plotting the linear

* Dr Göksu Ögelman has been dismissed from her post by the Martial Law Authorities in Turkey (see *Nature* **295**, 638).

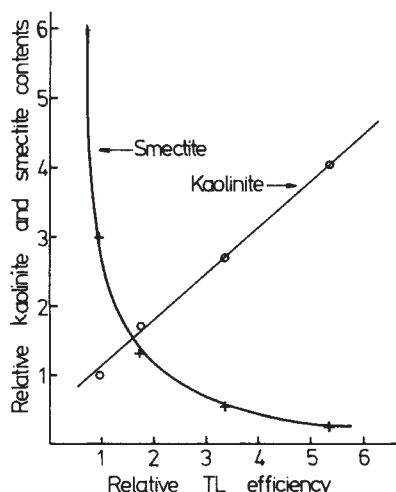


Fig. 2 Variation of relative kaolinite and smectite content with TL efficiency.

portion of the graphs of radiation dose against the area under the 275 °C peak; the saturation dose levels were determined from the asymptote of the curves.

The products from weathering of the same samples of feldspars were observed by polarizing microscope and X-ray diffraction techniques. The relative kaolinite and smectite contents were determined by measuring the area under the 7.2 Å and 18.0 Å peaks respectively, of powdered X-ray diffractograms.

Figure 2 summarizes the relative behaviour of kaolinite, smectite contents of weathered feldspars with respect to TL efficiency. As weathering progresses the kaolinite content increases, the smectite content decreases, and the TL efficiency and saturation dose level increases.

Figure 3 shows a thin section of the basalt from profile 2, covered with feldspars. Weathering can be observed as the cracking of the twinning planes as well as the formation of weathered regions (clay minerals) on the surface. Therefore, it is unlikely that improvement of optical transmission properties could account for the increased TL efficiency. Because the TL efficiency is proportional to the density of lattice defects, it is likely that weathering increases the defects on the surface of feldspars. Scanning and transmission electron microscopy could help to estimate the type of defects produced during the transformation of feldspars to clay with the TL method.

TL is a quick and cheap method compared with the usual lengthy mineralogical procedures for determining the relative

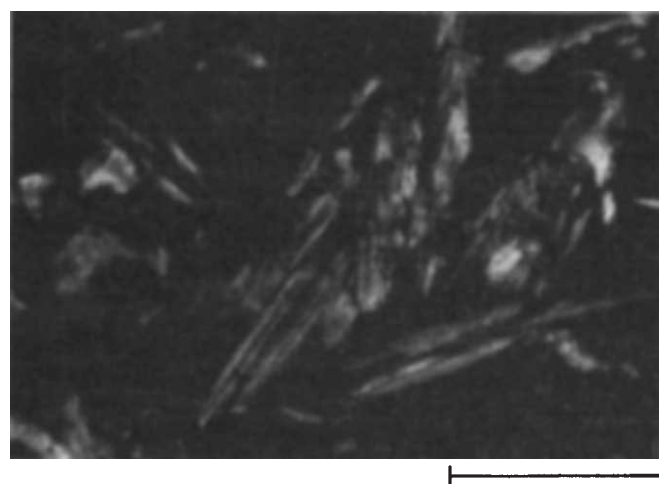


Fig. 3 A thin section of one of the basalt rocks with moderately weathered feldspar grains from profile 2. Weathering can be observed as the cracking of their twinning planes as well as the formation of weathered regions (clay minerals) on the surface.

stages of weathering in basaltic rocks. Conversely, the TL characteristics of weathered basaltic rocks in the same regions or similar climatic and chemical environments, can be used to determine their relative ages. These concepts could be extended to other weathered rock materials.

We thank the Turkish Scientific and Technical Research Council for supporting this research through the Archaeometry Study Group. Y.G.Ö. further thanks her fellow prisoners for friendship and understanding, and her husband whose continued love and support eased the agony of life in prison and made the completion of this article possible.

Received 5 January; accepted 22 January 1982.

1. Kapur, S., Özbek, H. & Dinç, U. *Proc. Vth Congr. of the Turkish Scientific and Technical Research Council*, 87–100 (Tübitak, Ankara, 1977).
2. Singer, A. & Navrot, J. *J. Soil Sci. Soc. Am.* **41**, 645–650 (1977).
3. Aitken, M. J., Zimmerman, D. W. & Fleming, S. J. *Nature* **219**, 442–445 (1968).
4. Fleming, S. J. *Archaeometry* **12**, 135–146 (1970).
5. Zimmerman, D. W. *Archaeometry* **13**, 29–52 (1971).
6. Göksu, H. Y., Fremlin, J. H., Irwin, H. T. & Fryxell, R. *Science* **183**, 651–654 (1974).
7. Aitken, M. J., Fleming, S. J., Doell, R. R. & Tanguy, J. C. *Thermoluminescence of Geological Materials* (ed. McDougall, D. J.) 359–366 (Academic, New York, 1968).
8. Wintle, A. G. & Huntley, D. J. *Nature* **279**, 710–712 (1979).
9. Christodoulides, C. & Fremlin, J. H. *Nature* **232**, 257–258 (1971).
10. Wintle, A. G. *Nature* **245**, 143–144 (1973).

A premonitory pattern of earthquakes in northern Greece

B. C. Papazachos, T. M. Tsapanos
& D. G. Panagiotopoulos

Geophysical Laboratory, Aristotelian University of Thessaloniki,
Thessaloniki, Greece

On 20 June 1978 at 20 h 03 min 23 s, an earthquake of surface wave magnitude 6.5 occurred in the Migdonian basin, between the Lagada and Volvi lakes, and 26 km from Thessaloniki. The main shock was preceded by a seismic sequence which started on 8 May. The largest earthquake of that sequence occurred on 23 May and had a surface wave magnitude of 5.8. The long duration of this sequence before the main shock of 20 June had tremendous psychological effects on the whole country. Greek seismologists were in a difficult position during that time (8 May–20 June), especially after the earthquake of 23 May, when pressed to state whether these earthquakes were foreshocks or aftershocks. Long-lasting seismic sequences with earthquakes of magnitude <6 are common in the Aegean area. Because these sequences are sometimes followed by large earthquakes it is important to see whether the latter can be predicted. The May–June 1978 seismic sequence provided the opportunity for testing hypotheses about statistical properties which, it has been suggested, could distinguish between foreshocks and aftershocks. This particular sequence has been investigated extensively^{1–4} and more accurate data are available than for any other sequence around Greece. Focal mechanism studies^{5,6} and field observations² showed that the largest earthquakes of the sequence were produced by horizontal extension in a nearly north–south direction and that the main rupture occurred on a normal fault which strikes almost east–west and dips to the north. The fault length of the main shock of 20 June was estimated⁷ to be 32 km. Accurate determination of earthquake foci, using portable seismographs, showed that the earthquakes of this sequence were shallow ($h < 17$ km) while the distribution of the epicentres had a very clear east–west elongation^{4,5}. A migration of the seismic activity from east to west has also been observed between 8 May and 14 July. We investigate here the time variation of the average magnitude and the frequency of all the shocks with $M \geq 3.0$ of this sequence. The time variations of these seismic parameters are examined in connection with the distribution of well determined epicentres of 38 earthquakes with $M \geq 4.3$ in order to search for any premonitory pattern of the 23 May and 20 June major earthquakes.

Accurate epicentre and focal depth determination for large and small earthquakes of the sequence exist for the period after 3 July 1978 when the first network of portable stations was put in operation in the epicentral region by D. Carver and R. Henrisey of the US Geological Survey. Two other such networks were put in operation, one on 20 July by a team under G. King of Cambridge University in cooperation with the Geophysical Laboratory of the University of Thessaloniki, and another on 1 August by the Public Power Corporation of Greece (PPC). However, one portable seismograph run by the Geophysical Laboratory was in operation in Thessaloniki for the period 26 June–2 July.

For the period 8 May–2 July, the available data come from the network of permanent stations of Greece and neighbouring countries and from a network of three permanent stations of PPC. The closest of these stations are PLG (40.37°N, 23.45°E) at a mean distance of 40 km south of the epicentral region, VAY (41.32°N, 22.57°E) at a mean distance of 90 km north-west of the epicentral region and the network of the three stations of PPC (KAP: 40.36°N, 21.98°E, POL: 40.31°N, 22.11°E, SER: 40.18°N, 22.01°E) at a mean distance of ~110 km south-west of the epicentral region.

As there was continuous recording from the three PPC stations during the whole sequence, their records were the basis for the present investigation of the time distribution of the earthquakes between 8 May and 2 July. The magnitudes of all these earthquakes were determined using relations of the form $M = a \log \delta t + V\Delta + c$, where, Δ is the epicentral distance⁸ and δt is the signal duration⁸. The constants a , V and c were determined by the usual statistical procedure for earthquakes for which M was known by other means. These formulae were used to determine a magnitude for each station and the average of the three determinations was taken as the magnitude of each earthquake. Magnitudes down to 2.6 were found. In attempting to determine the epicentres for all the earthquakes, we found that for those with $M \geq 2.8$ we could decide whether their epicentres were in the epicentral region of the sequence or elsewhere. However, we finally decided to include in our catalogue and use in this study only earthquakes with $M \geq 3.0$. In this way we have a catalogue of complete and homogeneous data with respect to magnitude for the whole period of the sequence. Between 8 May and 2 July, the number of shocks with $M \geq 3.0$ was 393. This catalogue will be published elsewhere.

The statistical relationship $\log N = a - bM$ between the number, N , of earthquakes which have magnitudes equal to or larger than a certain value and this value, M , holds also for seismic sequences (foreshocks, aftershocks, swarms). The parameter b in the frequency–magnitude relationship is smaller for foreshocks than for aftershocks^{9,10}; laboratory experiments have shown that this can be attributed to the homogeneity of the material¹¹ or to the high stress level¹² in the focal region during the foreshock period. That is why the low values of b have been suggested as precursors of strong earthquakes¹³. However, it is not easy to determine b accurately for several time intervals to follow its pattern, thus the time variation of other parameters, which are related to b but are easily and accurately determined, can be examined to see whether they establish any premonitory patterns of strong earthquakes. The difference in magnitude between the main shock and the largest aftershock has been suggested¹⁴ as such a parameter.

We now present evidence that the time variation of the average magnitude value could be considered as premonitory pattern of the 23 May and 20 June 1978 earthquakes. Relationships between $\bar{M}$ and the parameter b have been derived by Hamilton¹⁵ and Utsu¹⁶. Utsu gave the following simple formula:

$$\bar{M} = \frac{\log e}{b} + M_{\min} \quad (1)$$

where $\log e = 0.4343$ and $M_{\min}$ is the magnitude of the smallest earthquake considered. This relation shows that $\bar{M}$ must have relatively large values for foreshocks and smaller values for aftershocks.

The time variation of the average magnitude of this sequence was determined as follows: a minimum number k of events was defined and, for each day, the average of all earthquake magnitudes, with $M \geq M_{\min}$ was taken as the mean magnitude $\bar{M}$ when the number of events n of that day was $\geq k$. When a day had a total number of considered events $n < k$ then the $k - n$ events were taken from its previous day (or days). In this way, each value of $\bar{M}$ was determined with k or more observations.

Such calculations of $\bar{M}$ were made with all earthquakes in the catalogue ($M_{\min} = 3.0$) and $k = 10$. The calculations were repeated with $k = 15$. Plots of $\bar{M}$ against time have shown that the whole seismic sequence can be separated in four sub-sequences. These sub-sequences cover the periods 8 May–23 May–1–20 June–2 July, during which the $\bar{M}$ values are 3.67, 3.35, 3.61, 3.48 and the corresponding number of shocks with $M \geq 3.0$ are 24, 57, 65, 247, respectively. This means that the average values of magnitudes are relatively high–low–high–low during these four consecutive time periods. Unfortunately no data exist to determine the average magnitude for the background seismic activity of the area. These data will become available in the next few years because the Geophysical Laboratory installed a telemetry network of eight seismic stations in the area on 1 January 1981. The calculations were repeated with other values of $M_{\min}$ (3.1, 3.2). The decrease in the sample number had some effect on the plots but the basic pattern remained unchanged.

Figure 1a shows the time variation of $\bar{M}$ for $k = 15$ and $M_{\min} = 3.0$. The values for the period 8–13 May have not been plotted because there are not enough data ($n < 15$) to determine $\bar{M}$ with the accuracy of the other plotted values, but the values

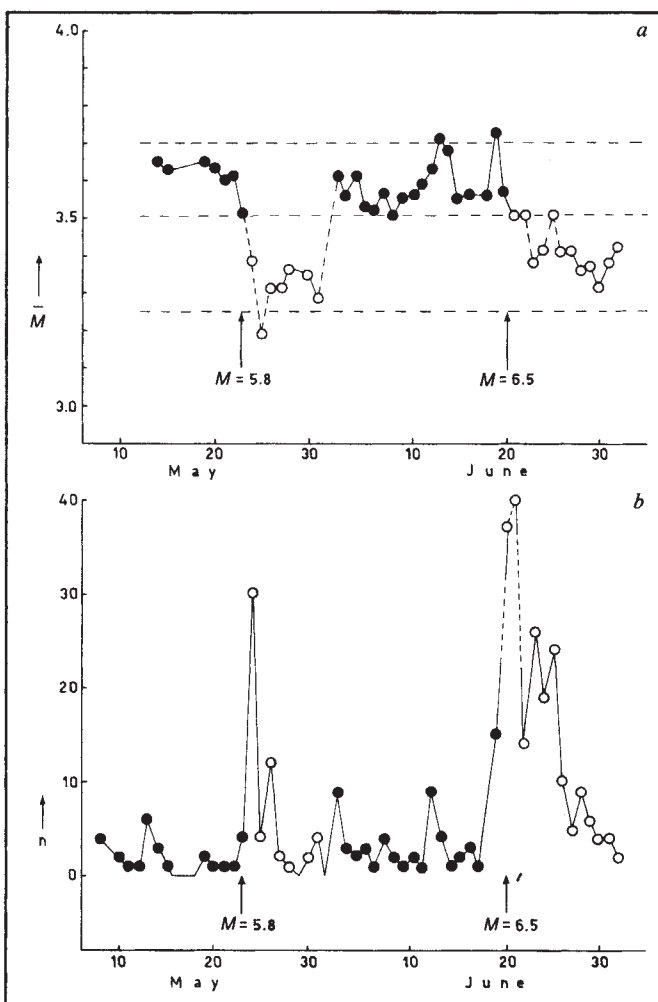


Fig. 1 Time variation of the mean magnitude (a) and of the daily number (b) of the earthquakes of the sequence with $M \geq 3.0$. ●, Foreshocks; ○, aftershocks.

Table 1 Information on all earthquakes with $M \geq 4.3$ of the sequence

Date	Origin time	φ_N	λ_E	h	M	N	s.e.	Ref.
8 May 1978	14:38:58	4040.6	2320.5	14	4.7	40	1	4
	15:00:08	4044.4	2323.1	10	4.3	14	4	*
10 May 1978	13:12:51	4041.5	2322.4	17	4.6	25	3	*
13 May 1978	08:35:36	4040.6	2324.0	11	4.5	23	3	*
19 May 1978	14:46:08	4041.1	2323.5	15	4.4	28	2	*
23 May 1978	23:34:13	4041.9	2317.7	8	5.8	240	1	4
	02:12:30	4039.2	2323.1	14	4.7	84	2	4
24 May 1978	05:57:28	4041.9	2318.3	13	4.6	54	1	4
	08:13:09	4043.9	2322.9	—	4.4	24	4	*
	08:46:31	4042.4	2319.5	—	4.4	21	3	*
2 June 1978	22:31:25	4042.9	2314.6	14	4.9	115	2	4
	22:40:13	4044.6	2319.3	6	4.3	25	3	*
12 June 1978	17:44:48	4045.4	2322.4	11	4.5	37	2	*
	23:36:42	4045.9	2317.2	16	4.5	30	2	*
13 June 1978	01:36:54	4042.7	2317.6	5	4.3	23	2	*
	03:12:54	4042.1	2319.6	11	4.5	21	3	*
19 June 1978	10:31:07	4042.4	2315.5	12	5.2	205	1	4
	10:48:12	4040.9	2314.8	11	4.6	87	1	4
	20:03:23	4043.7	2315.2	8	6.5	385		4
20 June 1978	20:37:41	4041.8	2309.6	—	4.4	19	4	*
	20:45:25	4048.5	2314.7	—	4.5	22	3	*
	20:52:42	4045.8	2309.0	17	4.3	20	3	*
	21:51:06	4043.6	2314.9	—	4.6	28	2	*
	03:20:27	4044.9	2312.4	15	4.6	23	2	*
21 June 1978	06:00:18	4039.0	2316.2	—	4.4	48	3	4
	12:29:46	4044.0	2305.9	0	4.7	94	2	4
	18:52:06	4041.2	2312.9	10	4.4	58	1	4
	21:20:14	4043.2	2310.0	13	4.3	20	4	*
22 June 1978	02:26:42	4045.7	2309.8	—	4.3	20	4	4
23 June 1978	01:57:03	4043.0	2304.7	2	4.5	30	3	4
26 June 1978	00:04:01	4046.3	2309.1	—	4.5	27	2	*
4 July 1978	22:23:28	4042.0	2306.4	8	5.1	148	2	3
13 July 1978	17:26:57	4040.9	2306.1	6	4.4	38	2	3
15 July 1978	01:42:03	4043.8	2318.1	4	4.3	41	1	*
21 August 1978	00:52:46	4042.0	2330.1	6	4.3	7	4	4
24 August 1978	01:23:51	4041.5	2326.7	4	4.3	7	2	4
11 May 1979	01:46:28	4044.0	2321.8	9	4.7	18	1	*
31 August 1979	17:24:10	4042.8	2319.8	11	4.8	31	2	*

* Our determinations.

for this period are clearly high because only shocks with $M \geq 3.2$ occurred during this time. The time of occurrence of the two major earthquakes of 23 May ($M = 5.8$) and of 20 June ($M = 6.5$) are also shown in Fig. 1a, while in Fig. 1b the frequency (daily number) of earthquakes with $M \geq 3.0$ as a function of time is shown.

From Fig. 1 we see that during the two aftershock periods (23 May–1 June, 20 June–2 July), when the mean frequency of earthquakes decreases with time, the average magnitude values are clearly low. The mean of the average magnitudes of these periods is 3.38 with a s.d. of 0.06. During both the 15 days before the earthquake of 23 May and during the 19 days before the earthquake of 20 June, the mean magnitude values are clearly high. The mean of the average magnitudes of these two periods is 3.60 with a s.d. of 0.08. It seems, therefore, that the 23 May earthquake was preceded by its foreshocks and followed by its own aftershocks. The latter were distinct from the foreshocks of the 20 June earthquake. The standard deviations above mentioned have been multiplied by 1.6 to determine approximately the 90% confidence intervals. The three horizontal dashed lines in Fig. 1a show these 90% confidence intervals. The upper limit of this interval for the aftershocks coincides with the lower limit of this interval for the foreshock period.

The high values of $\bar{M}$ during the whole seismic period before the earthquake of 23 May as well as during the 19 days period before the earthquake of 20 June could be considered as precursors of these two earthquakes. Similar observations in the future could possibly be used as a warning signal.

The frequency of aftershocks after the main shock of 20 June reached its first minimum on 2 July and therefore the earthquakes which occurred during this period can be considered as

real aftershocks. That is why we considered here only shocks which occurred up to 2 July. The number of shocks with $M \geq 3.0$ for the periods 3–31 July, 1–31 August and 1 September–31 December 1978 was 19, 26 and 31 respectively. That is, the number of shocks are less than one event per day even for July and August. So, it is impossible to determine reliable daily values of $\bar{M}$ and plot them. Instead, the mean magnitudes for each of these periods have been determined and values of 3.49, 3.49 and 3.43 have been found. These values are within the 90% confidence intervals for aftershocks but are relatively high, especially the first two. This is probably due to the expansion of the activity to the west during July and to the east during August in areas which were not previously affected and some earthquakes of these periods considered as aftershocks of the 20 June event were probably foreshocks of second order sequences.

Unfortunately, no network of portable seismographs was operating in the epicentral area before 3 July with which to investigate the space–time distribution of a large number of earthquakes during 8 May–2 June. However, some of the largest earthquakes of this period have been determined accurately^{3,4} using the joint epicentre method. As these determinations are not sufficient to show the space–time distribution of the earthquakes which occurred before 3 July 1978, we have attempted to determine as accurately as possible the epicentres of all earthquakes of the sequence with magnitudes larger than a certain value. For this purpose the Hypo 71 program has been used with a three layer crustal model which is usually applied to determine epicentres in the Aegean area. To increase the accuracy, time delays determined from data of earthquakes recorded by portable seismographs (after 2 July) were added to the arrival times at the permanent stations in Greece and

neighbouring countries. It was found that the epicentres of all shocks of the sequence with $M \geq 4.3$ were determined with standard errors ≤ 4 km. On the other hand, comparison of 20 epicentre determinations made by this method and the determinations of the same epicentres by other methods and data (joint epicentre, portable seismographs) showed very good agreement.

Table 1 gives information for all earthquakes of the sequence with $M \geq 4.3$. The number of observations, N , used to determine the focal coordinates, and the standard error, s.e., of the epicentre are given. All these 38 epicentres are plotted in Fig. 2.

The epicentre of the earthquake of 23 May as well as the epicentres of its largest ($M \geq 4.3$) foreshocks and aftershocks are shown in Fig. 2a. The epicentres of these earthquakes (8 May–1 June) are located in the eastern part of the seismogenic volume of the entire sequence, near the western coast of the Volvi lake.

Figure 2b shows the epicentre of the main shock of 20 June, the epicentres of the largest ($M \geq 4.3$) earthquakes which occurred during the 19 days period which preceded this earthquake (1–20 June) and the epicentres of the largest ($M \geq 4.3$) earthquakes of the period 20 June–14 July. The largest earthquakes of the period 1–20 June had epicentres in a volume close to the epicentre of the main shock north of the area between the two lakes. This supports the idea that these earthquakes are real foreshocks of the earthquake of 20 June. Note that the largest two of these foreshocks occurred near the beginning (2 June, $M = 4.9$) and the end (19 June, $M = 5.2$) of this period and were located very close to the epicentre of the main shock. The largest aftershocks of the 20 June earthquake

which occurred for 23 days after the main shock were located further west in a volume clearly different from the foreshock volume. Such a westward migration of seismic energy release is also clearly shown in the distribution of the macroseismic intensities.

Figure 2c shows the epicentre distribution for the largest ($M \geq 4.3$) earthquakes which occurred after 14 July 1978. Note that the main activity during this period shifted to the eastern part of the area.

The distribution of epicentres shows clearly that between 23 May and 1 June only the eastern part of the seismic volume was active. During this period as well as during the rest of the time until the occurrence of the main shock, the central part of the area, where this main shock of 20 June was located, was in very high stress condition. Because the main seismic activity of the period 1–20 June occurred in this central part of the seismic volume in high stress conditions, high values of $\bar{M}$ (low values of b) are expected during this period. Therefore, the high $\bar{M}$ values (or low b) as well as the space–time distribution of the earthquakes which occurred between 1 and 20 June show that these are real foreshocks. If this information had been known in June before the main event ($M = 6.5$) it would have helped seismologists to predict that the danger was not then over.

We thank the Public Power Corporation of Greece for use of the records from its stations.

Received 8 October; accepted 31 December 1981.

1. Papazachos, B. C., Moundrakis, D., Psilovikos, A. & Leventakis, G. *Tectonophysics* **53**, 171–183 (1979).
2. Mercier, J. L., Mouyaris, N., Simeakis, C., Roundojaynis, T. & Angelindhis, C. *Nature* **278**, 45–48 (1979).
3. Carver, D. & Bollinger, G. A. *Tectonophysics* **73**, 343–363 (1981).
4. Soufleris, C., Jackson, J. A., King, G. C. & Spencer, C. P. *Geophys. J. R. astr. Soc.* (in the press).
5. Papazachos, B. C., Moundrakis, D., Psilovikos, A. & Leventakis, G. *Bulgar. geophys. J.* **6**, 72–80 (1978).
6. Soufleris, C. & Stewart, G. S. *Geophys. J. R. astr. Soc.* **67**, 343–358 (1982).
7. Kulhanek, O. & Meyer, K. *Special Volume on the Thessaloniki 1978 Earthquake* (in the press).
8. Lee, W. H. K., Bennett, R. E. & Meagher, K. L. *U.S. Geol. Surv. Open file Rep.* (1972).
9. Suyehiro, S. *Bull. seism. Soc. Am.* **56**, 185–200 (1966).
10. Papazachos, B. C., Delibasis, N., Liapis, N. & Moumoulidis, G. *Ann. Geofys.* **20**, 1–93 (1967).
11. Mogi, K. *Bull. Earthq. Res. Inst. Tokyo* **41**, 595–614 (1963).
12. Scholz, C. H. *Bull. seism. Soc. Am.* **58**, 399–415 (1968).
13. Scholz, C. H., Sykes, L. R. & Aggarwal, Y. P. *Science* **181**, 803–809 (1973).
14. Papazachos, B. C. & Comninakis, P. E. *Tectonophysics* (in the press).
15. Hamilton, R. H. *Bull. seism. Soc. Am.* **58**, 399–415 (1967).
16. Utsu, T. *Bull. Hokkaido Univ.* **13**, 99–103 (1965).

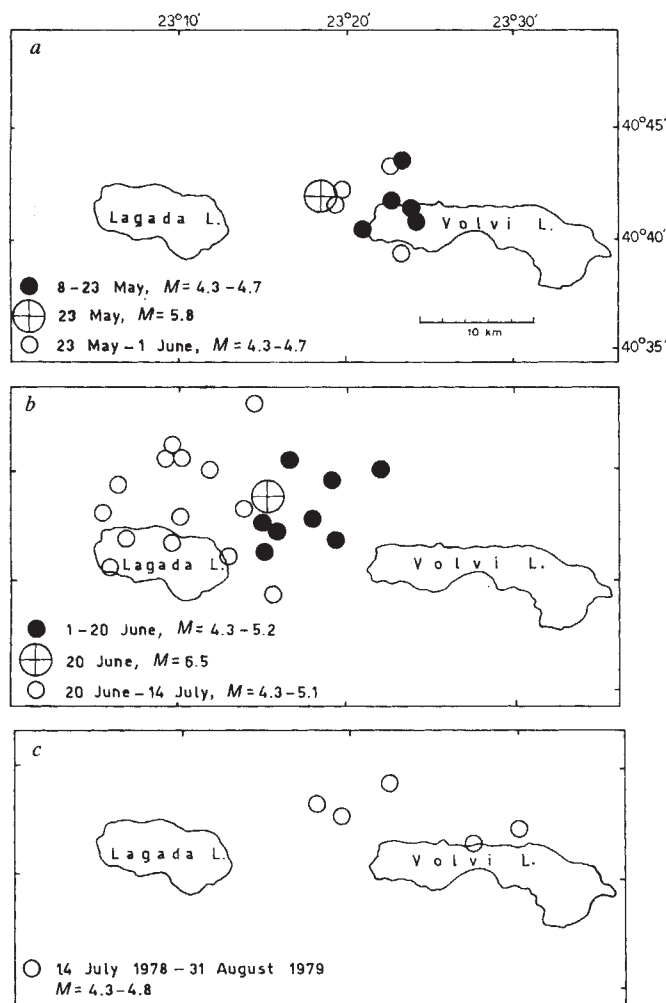


Fig. 2 Epicentres of all earthquakes of the sequence with $M \geq 4.3$ for the periods 8 May–1 June, 1978 (a), 1 June–14 July, 1978 (b) and 14 July, 1978–31 August 1979 (c).

Direct evidence of a subducted plate under southern Mexico

C. Lomnitz

Instituto de Investigaciones Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, México 20, D.F.

The hypothesis¹ of subduction of oceanic plates under active continental margins has become a fundamental idea in plate tectonics. Yet the evidence for the existence of oceanic plates beneath continental borders is still largely circumstantial. In seismology, the presence of a new layer or structural unit is taken as confirmed only after seismic phases corresponding to reflections or refractions from that layer have been identified on seismograms. I report here the presence of a mantle phase, interpreted as a seismic wave refracted from a dipping interface on records of aftershocks from the intermediate earthquake of 24 October 1980 in the Mixtec Highlands of south-central Mexico which may be direct evidence on the structural position of the subducted plate under the Mexican active continental margin.

The destructive earthquake of 24 October 1980 in the Mixtec Highlands ($M_L = 7.4$, $h = 70$ km) generated over 50 aftershocks

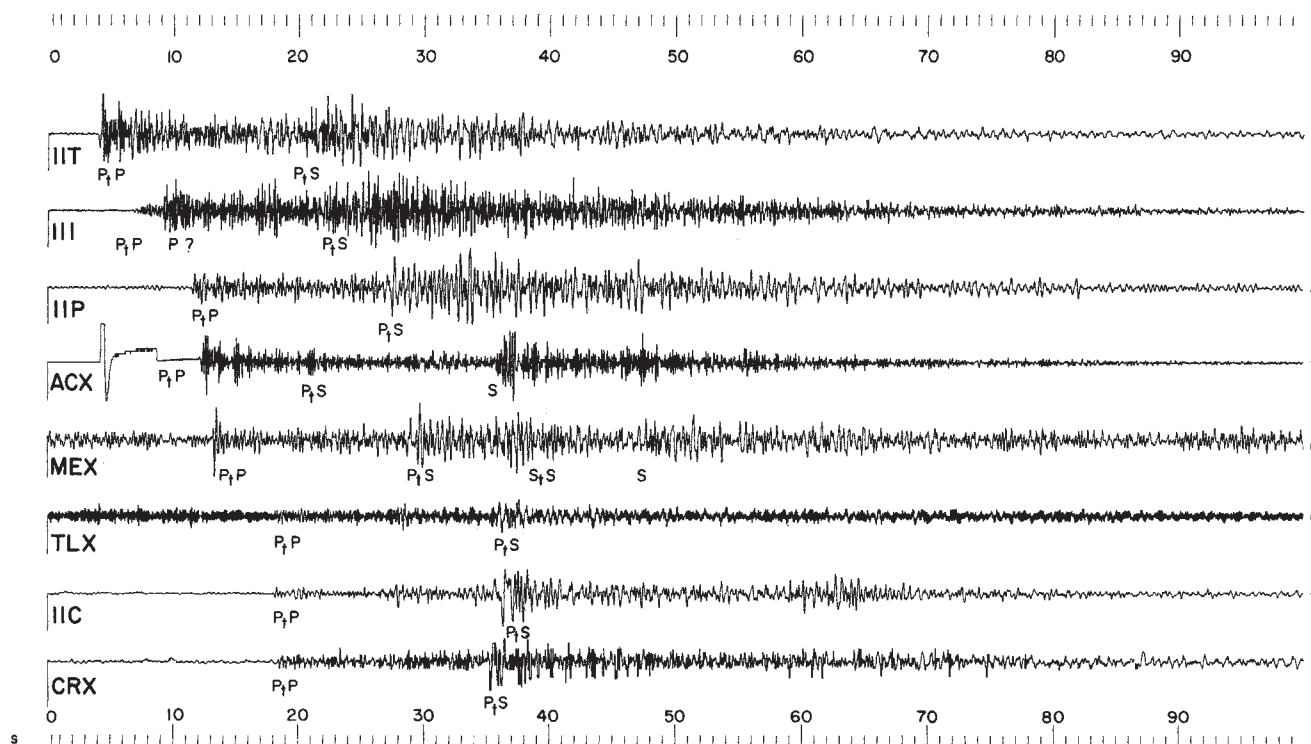


Fig. 1 Seismograms of aftershock 15:076° of the RESMAC catalogue, showing plate phases. All records are from 1-Hz vertical instruments (time in seconds).

which were large enough to be recorded and located on the RESMAC telemetric network. On each record a prominent arrival is found between 10 and 20 s after the initial P signal (Fig. 1). This phase cannot be identified as S, in spite of its low-frequency content, because S is expected to arrive (and is actually seen on most seismograms) 10 or more seconds later. It is not a P-to-S conversion at the bottom of the crust, as this would require a crustal thickness of >100 km and a refraction profile measured in this region in 1975 reveals that the crust is ~ 40 km thick² (S. K. Singh *et al.*, in preparation).

The possibility that the mantle phase, which we call P₁S, might represent a crustal phase must first be discarded. The focal depth

of 70 km reported by Boulder on the strength of depth phases has been questioned²; however, independent evidence on the focal depth of the aftershocks exists from the records of a four-station temporary network operated in the epicentral region by the Institute of Geophysics of the National University of Mexico (R. Mota, in preparation). The S-P times at the nearest station are typically 7–10 s, indicating (if 40 km is the correct crustal thickness) a range of 55–85 km in focal depths of the aftershocks. If the aftershock activity had been in the crust, S-P times of <6 s should have been observed. The character of the seismograms also indicates a focus below the crust, because of the sharp P-onsets and the absence of the crustal arrivals

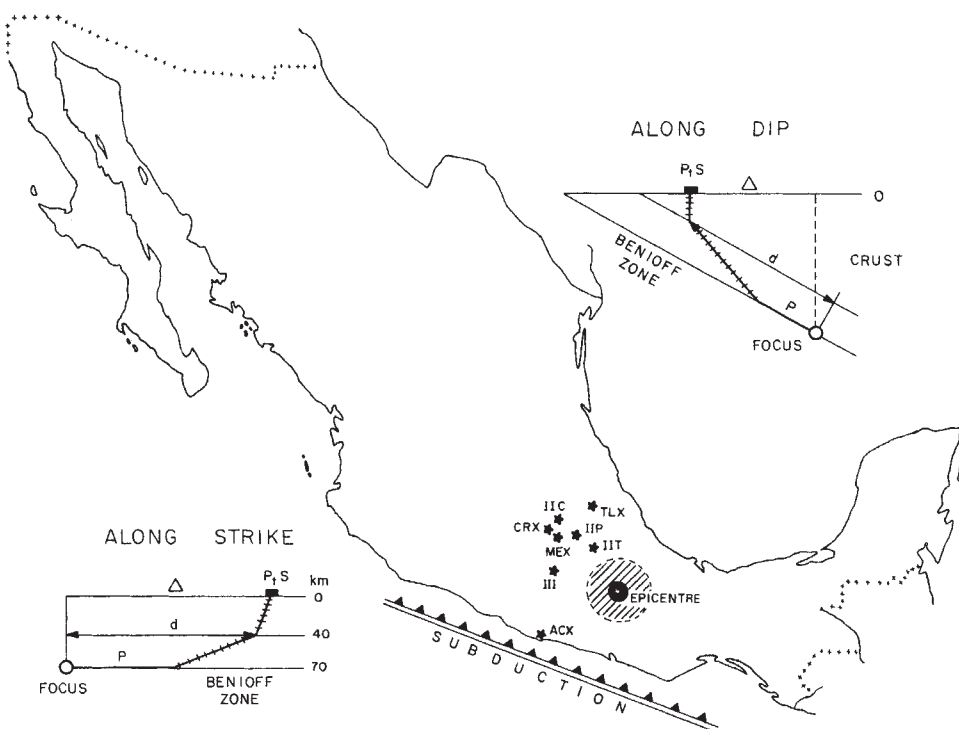


Fig. 2 Location map of RESMAC network in relation to the epicentral region, and proposed path of head wave P₁S refracted from the subducted Cocos Plate. Δ , epicentral distance; d , reduced distance for the same station.

which are typically found on seismograms from shallow events in this region.

The prominent amplitude and low-frequency content of the new phase suggest the presence of a refractor or waveguide at some depth below the crust. If the focus of the earthquake is within or close to the waveguide much of the seismic energy can be trapped and later refracted upwards as P or as S. This waveguide has exactly the structural position and velocity expected for the subducted Cocos Plate. Thus no plausible mechanism except subduction will account for the observations (Fig. 2).

Consider first the conventional travel-time plot constructed for the aftershock (Fig. 1). This event (25 October 1980 at 09:54:38UT) was relocated at 17.98° N, 98.12° W with the aid of the temporary stations. The travel times (Fig. 3) reveal that the apparent velocity of the second arrival is $\sim 8.0 \text{ km s}^{-1}$, too high for any direct or refracted second arrival from a layered crustal model. The expected arrival times for P_g and S_n are shown for comparison. Moreover, the converted P-to-S phase should arrive earlier, that is 6 s after the direct P arrival. Note that the fit of some individual arrivals is poor, notably for the coastal station ACX (Acapulco).

Assuming that the observed second arrivals represent a mantle phase, I present two different models for a deep refractor with a P-velocity of $8.2\text{--}8.5 \text{ km s}^{-1}$, corresponding to the bottom of the subducted oceanic crust (Fig. 4). First a simplified model is assumed involving two parallel dipping interfaces. The upper interface is the continental Moho and the lower interface is the oceanic Moho within the subducted plate (Fig. 4a). The geometry is such that the oceanic Moho crops out along the trench axis and the crustal thickness is 40 km at the epicentre. Moreover, the focus of the earthquake is assumed to lie at a depth of 70 km, precisely near the bottom of the subducted oceanic crust. The abscissae are distances measured along the dipping continental Moho, not along the surface of the Earth (see inset, Fig. 4a). In other words, the 'reduced distance' is the intercept of the refracted ray on the continental Moho, used as a datum plane. The Earth's surface cannot be used as datum because the refracted wave fronts describe a set of confocal ellipses on the surface: hence the plotting distance for a given wavefront varies with azimuth, and no rectilinear travel-time plot would be obtained.

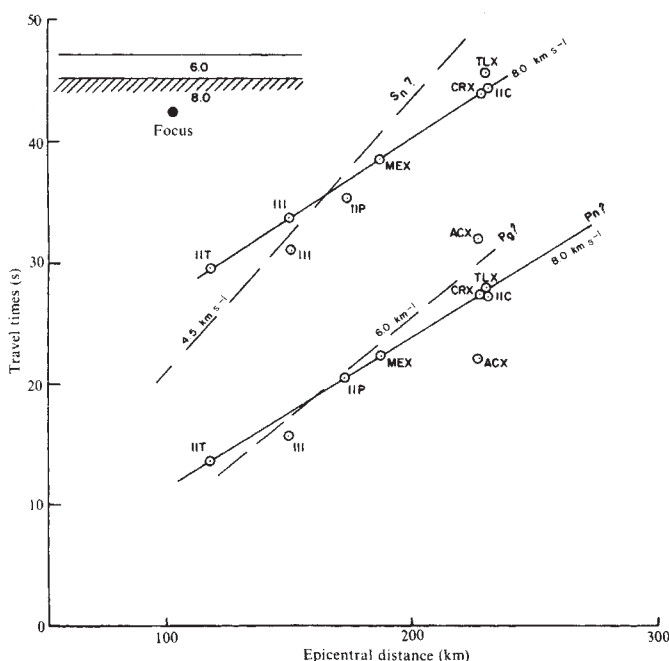


Fig. 3 Travel-time plot for a focus below the crust. Second arrivals do not fit any known phase. Note the poor fit of ACX.

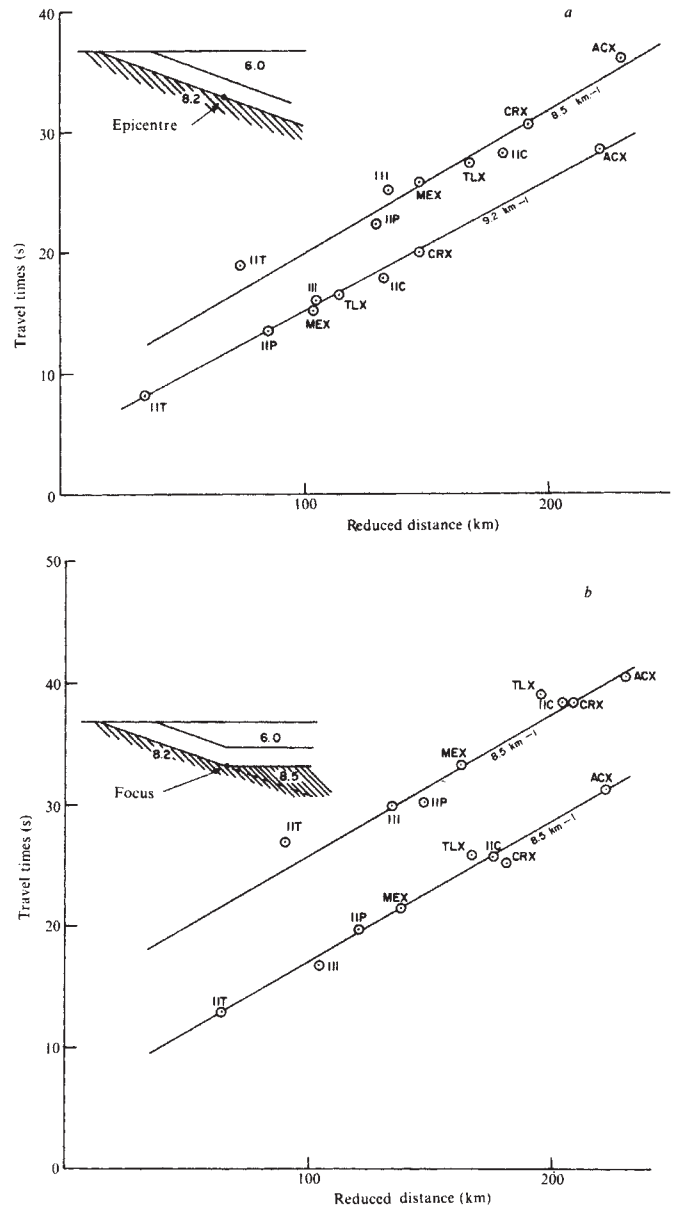


Fig. 4 a, A structural dip of 17° inland from the Mid-America Trench produces a good fit of the data, though the apparent velocities do not check with the model. b, A dipping subduction zone which levels off inland. The second arrival at III is doubtful (see Fig. 1), and might correspond to a direct S. Note that the apparent velocity of second arrivals is independent of the model.

A notable improvement of the travel-time fit is obtained by letting the deep refractor level off inland, as shown in Fig. 4b. This structural position was suggested by recent findings from refraction experiments on the coast of Colombia³. A slightly higher velocity for the inland position of the refractor ($v_p = 8.5 \text{ km s}^{-1}$) was assumed to conform with the Colombian results, though this is not an essential feature of the model. The apparent velocities of both refracted waves (P_g and P_n) are now equal and consistent with the model.

Mantle phases from intermediate and deep-focus earthquakes had previously been observed in Tonga, Japan and Peru⁴⁻⁸, though most of this work describes individual arrivals of phases other than P_g or P_n . The RESMAC array in central Mexico is well placed for observations of deep mantle phases; the apparent velocities of the phases can be determined and the approximate structural position of the deep refractor can be inferred. This proof of subduction is particularly convincing to the seismologist as it complements the structural information obtained from offshore explosions recorded on land.

The refractor is tentatively identified as the bottom of the subducted oceanic crust. The oceanic crust itself has a lower velocity, perhaps of the order of 6.8 km s^{-1} . If no such low-velocity lid existed the refractor would not function efficiently as a waveguide.

Consistently similar seismograms were recorded for over 50 aftershocks of the 1980 Mixtec earthquake. Similar deep refractions have also been observed on other Mexican intermediate-focus seismograms written on the RESMAC array. So-called 'early S-phases' recorded at local stations for the destructive Puebla-Veracruz earthquake of 28 August 1973 ($h = 84 \text{ km}$) can probably be interpreted as P_s. The present work demonstrates the potential of mantle phase observations for mapping the structural position of subducted plates under continental borders.

Does the subducted Cocos Plate level off inland, as indicated by the model of Fig. 4b? Earthquakes under the coastal plain of Veracruz, east of the Mixtec epicentre of 1980, typically have focal depths in the range of 80–100 km suggesting that the Benioff zone does indeed level off, because substantially greater depths would be expected for a uniformly dipping model as shown in Fig. 4a. Future studies of P_s observations from intermediate shocks in Veracruz should help resolve this question.

The RESMAC seismic network is supported by a grant of the Consejo Nacional de Ciencia y Tecnología of the Government of Mexico.

Received 21 July; accepted 30 October 1981.

1. Wilson, J. T. *Nature* **207**, 343–347 (1965).
2. Prince, J. (ed.) *Rep. IPS-8* (Instituto de Ingeniería, UNAM, Mexico, 1980).
3. Meyer, R. P. et al. *Geophys. Monogr.* No. 19, 105–132 (American Geophysical Union, 1976).
4. Mitronovas, W. & Isacks, B. L., *J. geophys. Res.* **76**, 7154–7180 (1971).
5. Isacks, B. L. & Barazangi, M. *Geophys. J. R. astr. Soc.* **33**, 129–139 (1973).
6. Snoke, J. A., Sacks, I. S. & Okada, H. *Phys. Earth. planet. Inter.* **9**, 199–206 (1974).
7. Sacks, I. S. & Snoke, J. A. *J. geophys. Res.* **82**, 2011–2017 (1977).
8. Nakanishi, I. *Tectonophysics*. **69**, 1–35 (1980).

Delineation of the North Central Italian upper mantle anomaly

G. F. Panza*, St. Mueller†, G. Calcagnile‡
& L. Knopoff§

* Istituto di Geodesia e Geofisica, Università di Trieste, Trieste, Italy
† ETH-Geophysics, Zurich, Switzerland

‡ Istituto di Geodesia e Geofisica, Università di Bari, Bari, Italy

§ Institute of Geophysics and Planetary Physics, University of California, Los Angeles 90024, USA

The Italian peninsula is the focus of intense deformational tectonic activity. It is underlain by mantle material whose inferred, relatively anomalous properties^{1,2} are not inconsistent with the seismic, volcanic and thermal activity that is manifested in this region. The most effective geophysical tool for mapping the structure of the uppermost 200–300 km of the Earth is the observation and analysis of seismic surface-wave dispersion on a regional scale. Here we synthesize the interpretations of Rayleigh wave dispersion measurements made by several authors, each for different parts of North Central Italy^{3–8}, to delineate the lateral extent of the upper mantle anomaly in this region.

Over the past 20 years intensive efforts have been made to study details of the cross-section of the crust and upper mantle in the European-Mediterranean area. As a result of these studies, several zones were delineated^{1,2}, such as the Western Mediterranean, the Tyrrhenian Sea, the Adriatic promontory of the African plate and the Central European Rift System, each with differing mantle structure. In addition to these struc-

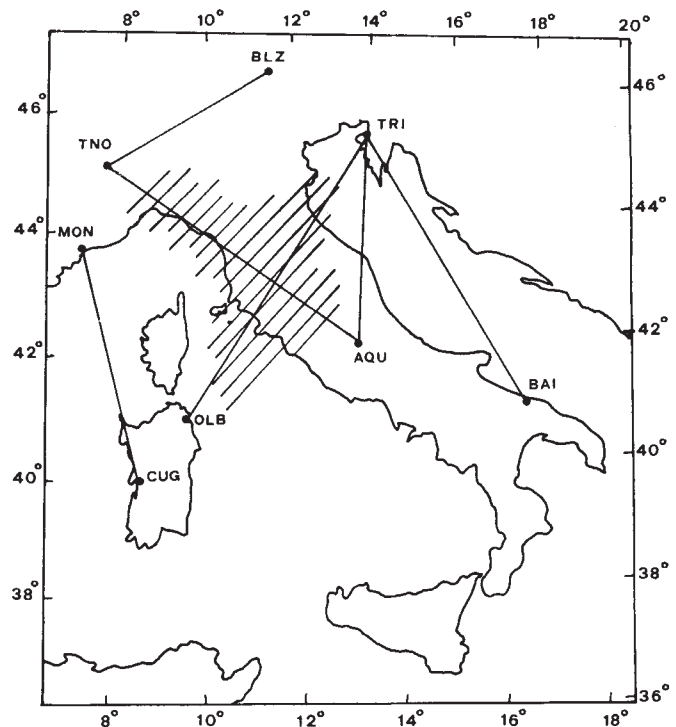


Fig. 1 Surface wave paths relevant to this study. The shaded region is the location of the mantle anomaly in North Central Italy.

tures, another conspicuous anomaly can be identified under the north-central part of the Apennine peninsula, that is the tectonically active region to the east and north of Corsica.

The properties of the upper mantle under North Central Italy (Fig. 1) have been delineated by Rayleigh wave dispersion measurements along two profiles TNO–AQU and TRI–OLB^{5,6}. These profiles have unusually low phase velocities in the period range 40–60 s (Fig. 2). Velocities as low as these have been observed in the Western USA⁷, as, for example, on the path TUC–BOZ, and the East African Rift¹⁰, both regions of high heat flow. The correlation between the surface wave phase velocities in northern Italy and high heat flow is even stronger after the phase velocities have been inverted; cross-sections are found consistent with the absence of a lid to the low-velocity channel (Fig. 3) and anomalously low S-wave velocities that extend up to, or close to the Mohorovicic discontinuity. In these cases the high surface heat flows are probably thus derived from high temperatures at shallow depth in the mantle and hence are a surface expression of tectonic involvement on a scale extending into the mantle.

The phase-velocity profile TNO–AQU (Fig. 2) differs in shape from the other two examples and would seem to provide an example of the danger of a literal interpretation of phase velocity curves in the period range 40–60 s. This curve differs from the others because of the influence of a larger crustal thickness under the axis of the Apennines than in the other regions described by low-phase velocities. Inversion of this phase velocity curve indeed indicates that a lid to a low S-wave velocity channel is absent⁵, and that this curve is properly catalogued with TRI–OLB, TUC–BOZ and others.

We can provide a rough limit to the lateral extent of the Italian anomaly. To the south, the region is abutted by the Tyrrhenian Sea and still farther south by the Sicilian crossing of the Mediterranean by the contact between the African and European plates. High heat flows, volcanism, low surface-wave phase velocities, deep focus earthquakes, and an inferred subduction¹¹ give the Tyrrhenian region the appearance of a back-arc basin. Thus the anomaly of the northern Italian peninsula seems to continue southward; the southern limit to the Tyrrhenian extension of the anomaly would appear to be at or

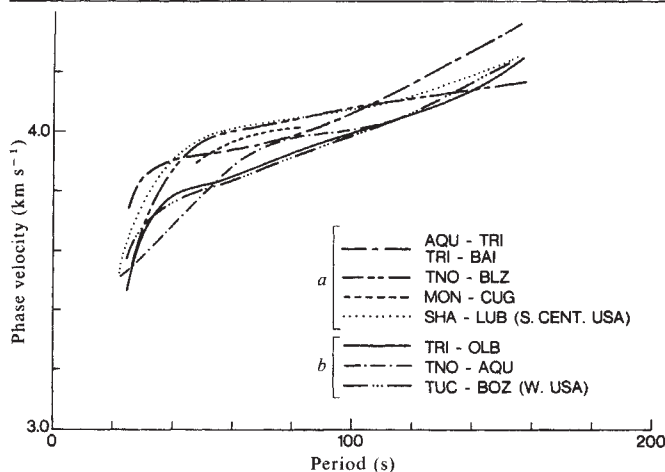


Fig. 2 Phase velocities used to delineate the North Central Italian anomaly and comparison with dispersion relations for western and South-Central USA. The r.m.s. errors for each curve are $\sim \pm 0.03 \text{ km s}^{-1}$ in the period range 40–60 s and are larger at the extremes of the period range^{3,5-9}. Curves in group *a* are for regions without tectonic involvement and in group *b* for regions with tectonic involvement.

near the plate boundary with North Africa. An exploration of this region will be described elsewhere.

To the west, surface-wave profiles near Corsica and Sardinia on path MON-CUG³, as well as to the north in the Italian Alpine foothills on path TNO-BLZ^{4,7}, have shown rather higher phase velocities in the period range characteristic of mantle depths. Similar phase velocities have been found elsewhere in relatively older, stable continental regions with more-or-less normal heat flows, that is, regions without major tectonic involvement today^{9,12,13} including the US Gulf Coast, such as on path SHA-LUB, and Western Europe. While inversion of these latter phase velocities may differ in some detail, especially in view of the non-uniqueness of the inversions, all successful inversions yield an upper mantle cross-section which has a broad low-velocity channel of significant contrast to an overlying high-velocity lid (Fig. 3); the channel-lid interface is found around a depth of 100 km. As indicated, this cross-section is evidently well correlated with the age and stability of these regions, characterized by their 'normal' heat flows and relatively

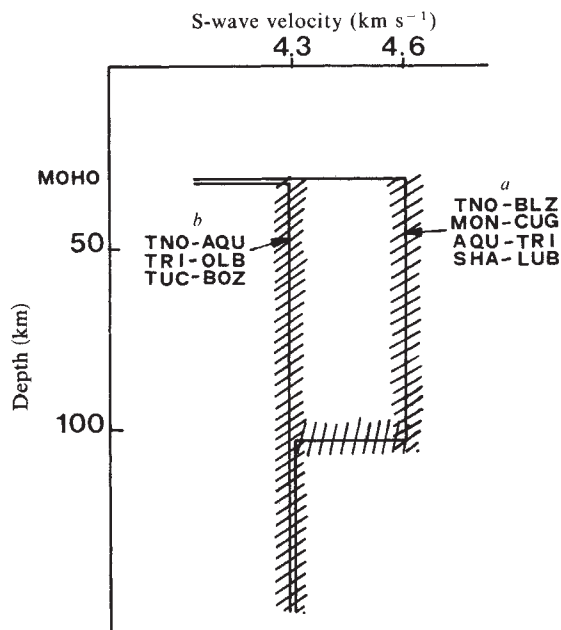


Fig. 3 Schematic shear wave velocity–depth distributions in the upper mantle for regions without *a* and with *b* major tectonic involvement today. The shading represents schematically the range of models satisfying each group of dispersion data.

low seismicity.

The values of phase velocities at periods between 40 and 60 s are diagnostic between the two groups. The errors in measurement are in every case less than the phase velocity interval between the groups of curves. It follows that the inversions of these phase velocity data yield structures for the two typical regions that are completely resolved from one another; the differences between the geophysical and geological environments of these two groups of regions are equally large.

The eastern edge of the Italian anomaly is delineated by surface-wave phase velocities that were determined along two Adriatic two-station paths, TRI-BAI and TRI-AQU⁸. The two phase velocity curves are indistinguishable over their common period range and are plotted as a single curve in Fig. 2. They are clearly members of the family of curves associated with young, stable continental regions, such as those for Western Europe. Such regions have well defined lids to channels. Cross-sections such as those for northern Italy, in which high-velocity lids are absent, are excluded. The inversion of these curves is no different, within errors of measurement, from those derived for other regions and sketched in Fig. 3.

We conclude that the anomalous low phase velocity region of northern Italy lies to the east of the path MON-CUG, to the west of the path AQU-TRI, and south of the Alps.

This work was supported by grants from CNR, Rome (80.2213) and NATO (1497), the Deutsche Forschungsgemeinschaft and the Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung.

Received 20 July 1981; accepted 19 January 1982.

1. Panza, G. F., Mueller, St. & Calcagnile, G. *Pure appl. Geophys.* **118**, 1209–1213 (1980).
2. Panza, G. F., Calcagnile, G., Scandone, P. & Mueller, St. *Le Scienze* May, 60–69 (1980).
3. Berry, M. J. & Knopoff, L. *J. geophys. Res.* **72**, 3613–3626 (1967).
4. Panza, G. F., Mueller, St. & Knopoff, L. *Rapp. Commun. int. mer. Medit.* **23**, 50 (1975).
5. Calcagnile, G., Panza, G. F. & Knopoff, L. *Tectonophysics* **56**, 51–63 (1979).
6. Sprecher, Ch. thesis, Swiss fed. Inst. Technol. (1976).
7. Calcagnile, G. & Panza, G. F. *Pure appl. Geophys.* **118**, 823–830 (1980).
8. Biswas, N. N. & Knopoff, L. *Geophys. J. R. astr. Soc.* **36**, 515–519 (1974).
9. Knopoff, L. & Schlue, J. W. *Tectonophysics* **15**, 157–163 (1972).
10. Caputo, M., Panza, G. F. & Postpischl, D. *J. geophys. Res.* **75**, 4919–4923 (1970).
11. Knopoff, L. *Tectonophysics* **13**, 497–515 (1972).
12. Seidi, D. & Mueller, St. *J. Geophys.* **42**, 283–328 (1977).

Pb isotopic composition of ophiolitic volcanogenic sulphide deposits, Troodos Complex, Cyprus

E. T. C. Spooner

Department of Geology, University of Toronto, Toronto, Ontario, Canada M5S 1A1

N. H. Gale

Department of Geology and Mineralogy, University of Oxford, Parks Road, Oxford OX1 3PR, UK

The volcanogenic sulphide ore deposits in the ophiolitic rocks of the Troodos Massif, Cyprus, are thought to be ~ 80 Myr old¹ analogues of the metal sulphide accumulations actively being deposited from $\sim 350^\circ\text{C}$ water at a depth of ~ 2.6 km on the East Pacific Rise (21°N)²⁻⁴. In both cases, mineralization is thought to be a consequence of heat transfer by convective seawater circulation within the oceanic crust^{5,6}. Evidence includes fluid inclusions in quartz from ore deposits in Cyprus with modal trapping temperatures of 298 – 370°C which have a bulk composition of 3.2 ± 0.5 equiv. wt % NaCl (1σ ; $n=273$). This salinity is identical to seawater with 3.5 wt% total dissolved solids^{4,7}. We have examined the isotopic composition of Pb as a geochemical tracer to clarify the origin of one of the metals

contained in the sulphide concentrations in the Troodos Complex. It was anticipated that the bulk of the lead would have been derived by leaching of the mafic ophiolitic rocks and, indeed, such a component is found to be present. However, the isotope ratios suggest that the lead is actually a mixture, and also contains an important component derived from seawater.

Interpretation of the isotopic data for the sulphides requires a knowledge of the isotopic characteristics of the associated igneous rocks of the ophiolitic complex. Hence, Pb isotope ratios were determined for a suite of eight variably ocean floor-metamorphosed basalts, dolerites and gabbros, the precise locations of which are given in Table 1 together with the results. Initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios were determined to monitor the degree of seawater-rock interaction which might also have affected the Pb isotopic composition, as this process is associated with increases in the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of altered rocks above initial magmatic $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of 0.70338 ± 0.00010 to 0.70365 ± 0.00005 (ref. 8).

Sulphide Pb isotope ratios were determined for 11 samples from five deposits (Table 2). In addition, a single independently obtained set of values is available for Skouriotissa⁹. As can be seen in Fig. 1, these six deposits give good spatial coverage of the Troodos Complex. In general terms, the sulphide concentrations consist of lenses of massive ore intercalated conformably within the pillow lava succession which are underlain by mineralized stockworks⁴. Massive ore is thought to have precipitated on the actual ocean floor from hydrothermal solutions which discharged through the stockworks beneath, and was deliberately avoided in this study because of possible complications associated with high level seawater mixing. Hence, most of the samples (10 out of 12) are from stockworks some of which are deep (for example, Alestos; ~1 km), others being relatively shallow (for example, Mathiati; ~40 m) relative to the original ocean floor. Two samples from the Limni mine in western Cyprus were mill concentrates and represent an average of the ore being processed at the time. All the other materials analysed were sulphide separates consisting mainly of pyrite and chalcopyrite produced from ~0.5 kg samples.

As shown in Fig. 2, most of the lead isotope compositions of the ophiolitic mafic rocks fall in, or near, the fields for mid-ocean ridge basalts¹⁰, or are less radiogenic. Some very low Pb isotope ratios were determined. For example, two gabbros (T74 and T137) have $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios of 16.620 and 17.050, 15.401 and 15.481, and, 36.511 and 36.801, respectively. A metadolerite (CY/75/80) has ratios of 17.398, 15.370 and 36.930. These values are all distinctly less radiogenic than any presented by Tatsumoto¹⁰. For some reason, two of the samples (CY/75/73 and 94B) seem to have slightly higher $^{207}\text{Pb}/^{204}\text{Pb}$ ratios than is typical (Fig. 2b). The Pb isotopic compositions of most of the samples are not likely to have been significantly affected by seawater interaction because sample CY/75/80, a metadolerite dyke with a contaminated initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.70505 ± 0.00003 , has unradiogenic Pb isotope ratios similar to fresh gabbros with initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of 0.70357 ± 0.00006 and 0.70362 ± 0.00006 , and a metabasalt sample (CY/75/75; no. 1 in Fig. 2) with a low initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.70342 ± 0.00003 has Pb isotope ratios which are near the radiogenic end of the defined range. In addition, the most Sr-isotopically contaminated sample (CY/75/94B; no. 8 in Fig. 2), with an initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.70585 ± 0.00003 , has initial Pb isotope ratios which are slightly less radiogenic than sample CY/75/75 (no. 1) discussed above. A possible exception is sample CY/75/71 (no. 7 in Fig. 2) with the most radiogenic Pb isotope ratios which plot between the mid-ocean ridge basalt fields and the seawater fields. This sample was notable because it contained a prominent hydrothermal epidote-sulphide vein which could have contained Pb dissolved in the circulating fluid of seawater origin. The Pb isotopic composition of the ophiolitic mafic rocks of the Troodos Complex is apparently comparable with that of mid-ocean ridge basalts, but also extends to lower values.

The Pb isotope ratios of the 11 sulphide separates from five ore deposits analysed in this study are more radiogenic than those for the ophiolitic mafic rocks (Table 2 and Fig. 2). The following ranges are defined: $^{206}\text{Pb}/^{204}\text{Pb} = 18.404\text{--}18.888$; $^{207}\text{Pb}/^{204}\text{Pb} = 15.551\text{--}15.634$; $^{208}\text{Pb}/^{204}\text{Pb} = 38.271\text{--}38.743$.

Table 1 Sr and Pb isotope ratios for mafic ophiolitic rocks from the Troodos Complex, Cyprus, listed in order of increasing $^{87}\text{Sr}/^{86}\text{Sr}$ ratio

Sample no.	Sample description	Sample locality	Rb (p.p.m.)	Sr (p.p.m.)	Rb/Sr	$^{87}\text{Sr}/^{86}\text{Sr}$	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
CY/75/75 (1)	Metabasalt with SiO ₂ amygdaloids	Aia Marina-Xyliatos	~1.3	88.8	~0.015	0.70342 ± 0.00003	18.246	15.495	37.941
T74(2)	Fresh gabbro	3 km east of Pedhoulas	—	—	—	0.70357 ± 0.00006	16.620	15.401	36.511
T137(3)	Fresh gabbro	200 m west of Troodhitis monastery	—	—	—	0.70362 ± 0.00006	17.050	15.481	36.801
CY/75/73 (4)	Metadolerite dyke	1 km west of Stavros tis Psokas	~0.7	~47	~0.016	0.70470 ± 0.00004	18.173	15.549	37.809
CY/75/94A (5)	Metagabbro	Agros	~0.3	82.5	~0.003	0.70475 ± 0.00005	18.207	15.480	37.648
CY/75/80 (6)	Metadolerite dyke	Xyliatos-Spilia	~0.4	~123	~0.004	0.70505 ± 0.00003	17.398	15.370	36.930
CY/75/71 (7)	Metadolerite dyke with epidote-sulphide vein	Lyso-Stavros tis Psokas	~0.5	~481	~0.001	0.70538 ± 0.00004	18.549	15.609	38.437
CY/75/94B (8)	Metadolerite dyke	Agros	~0.2	152	~0.001	0.70585 ± 0.00003	18.115	15.599	37.802

Rb and Sr were determined by X-ray fluorescence on a Philips 1410 instrument²⁰. Sr was extracted using standard ion exchange techniques and its isotopic ratio was analysed on a V.G. Micromass 30 mass spectrometer²¹. As all samples were found to have Rb/Sr ratios <0.02 age correction to 80 Myr BP, the age of the Troodos Complex¹, was unnecessary. Quoted errors are 2 s.d. and are ≤ 0.00005 . All results are quoted relative to an Eimer and Amend SrCO₃ standard isotopic ratio of 0.70800. The mean Eimer and Amend $^{87}\text{Sr}/^{86}\text{Sr}$ ratio measured in this work was 0.70807 ± 0.00002 . Pb was extracted using anodic electrodeposition²² and its isotopic composition analysed with a 12-inch radius mass spectrometer using a standard silica gel-phosphoric acid technique which yields ratios accurate to better than $\pm 0.1\%$. The values have not been age corrected to 80 Myr BP using independently determined U/Pb and Th/Pb ratios. Hence, initial ratios would be less radiogenic.

* From ref. 8.

Table 2 Pb concentrations and isotope ratios for sulphide separates from volcanogenic ore deposits, Troodos Complex, Cyprus

Sample no.	Sample locality	Locality description	Pb (p.p.m.)	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
CY/75/98A	Alestos	Deep stockwork in basal group	1.55	18.404	15.560	38.271
CY/75/98D	Alestos		3.3	18.480	15.563	38.367
—	Skouriotissa*	Massive ore	176	18.476	15.571	38.405
CY/75/91	Peristerka (New Kambia)	High level stockwork	1.7	18.526	15.564	38.441
CY/75/12	Limni	Medium level stockwork	12.9	18.525	15.616	38.518
CY/75/33	Limni	Medium level stockwork	26.9	18.543	15.595	38.420
CY/75/69	Limni	Medium level stockwork	39.0	18.546	15.634	38.599
CY/75/70	Limni	Pyrite concentrate				
		Chalcopyrite concentrate	ND	18.524	15.600	38.481
CY/75/87	Mathiati	High level stockwork	15.2	18.661	15.617	38.676
CY/75/89	Mathiati	Massive ore	30.4	18.658	15.599	38.651
CY/75/100B	Mousoulos	High level stockwork	3.6	18.887	15.587	38.743
CY/75/100C	Mousoulos	High level stockwork	5.6	18.888	15.551	38.612

Pb concentrations were determined by mass spectrometric isotope dilution with an accuracy of $\pm 2\%$. Either a solvent extraction step using methylisobutylketone (MIBK) or an HBr anion ion exchange column was used to remove Fe before extraction of lead by anodic electrodeposition²². Isotopic analysis was carried out by the method described in Table 1. Age correction to 80 Myr BP was unnecessary as sulphides contain very low levels of U and Th relative to lead.

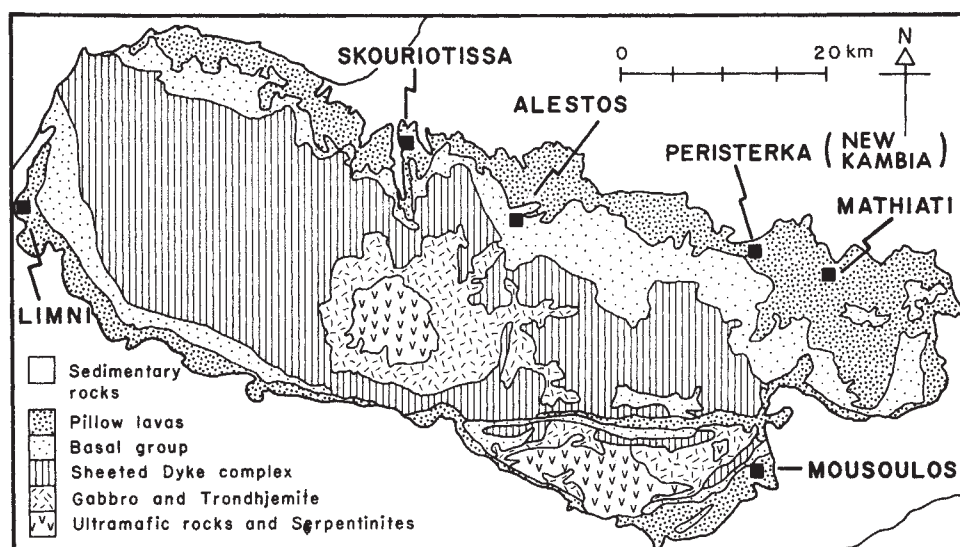
* From ref. 9.

Lead in the sulphide separates is present at trace levels of 1.55–39.0 p.p.m. averaging 14.0 ± 13.6 p.p.m. (1σ ; $n = 10$). These low concentrations reflect the lead-poor nature of ophiolitic volcanogenic sulphide deposits in contrast to Kuroko volcanogenic deposits which typically contain $\sim 1\%$ Pb, and are actually lower than lead concentrations of 115.0 ± 56.7 p.p.m. (1σ ; $n = 18$) determined by isotope dilution for Fe–Mn hydroxyoxide metalliferous sediments associated with ophiolitic rocks in Cyprus, Syria and The Sultanate of Oman¹¹. An independently determined set of ratios for the Skouriotissa deposit fall satisfactorily within the above ranges with corresponding values of 18.476, 15.571 and 38.405 (ref. 9). On $^{208}\text{Pb}/^{204}\text{Pb}$ against $^{206}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ against $^{206}\text{Pb}/^{204}\text{Pb}$ diagrams (Fig. 2), the ratios for the ophiolitic sulphide deposits in Cyprus define trends extending from near the mid-ocean ridge basalt fields (such as Alestos) to within the seawater fields defined by analysis of the Pb isotopic composition of manganese nodules^{12,13} (such as Mathiati). Values for three of the deposits (Skouriotissa, Peristerka and Limni) cluster in an intermediate position, whereas two sets of determinations for Mousoulos have the highest $^{206}\text{Pb}/^{204}\text{Pb}$ ratios but have slightly lower $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios than the sea-

water fields. The ranges enclose the values for the epidote-sulphide veined metadolerite dyke sample (CY/75/71; no. 7 in Fig. 2), which has already been noted as possibly Pb isotope contaminated, and are also surprisingly similar to the Pb isotope ranges for Fe–Mn hydroxyoxide sediments (umbers) associated with the ophiolitic rocks of Cyprus and the Baër-Bassit area in Syria¹¹. These oxidized metalliferous sediments are also thought to have formed as chemical precipitates from submarine hot springs. As the Pb isotope ratios for the sulphides occupy positions intermediate between the mid-ocean ridge basalt fields, which also contain values for the ophiolitic mafic rocks in Cyprus, and the seawater fields, the sulphide deposit lead apparently does contain a basaltic component, as expected, but the lead is actually a mixture which also contains lead derived from seawater.

The isotopic evidence for mixed lead in the ophiolitic sulphide deposits of Cyprus differs from data recently obtained for sulphide samples from 21°N on the East Pacific Rise^{14,15}. For example, Vidal and Clauer¹⁴ report $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios of ~ 18.48 , ~ 15.49 and ~ 37.90 . These values fall well within the MORB Pb isotope fields (Fig. 2) and correspond closely with the averages of the Pb isotopic

Fig. 1 Locations of the volcanogenic sulphide ore deposits in the ophiolitic rocks of the Troodos Massif, Cyprus, for which Pb isotope ratios have been determined in this study and by Doe and Zartman (Skouriotissa)⁹.



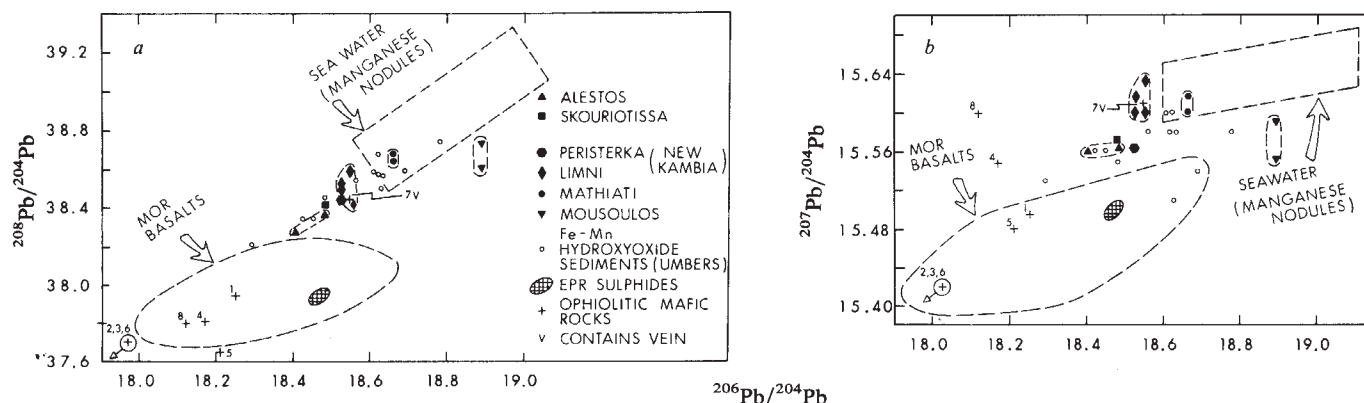


Fig. 2 Pb isotope ratios for volcanogenic sulphide ore deposits in the Troodos Complex, Cyprus, plotted together with ratios for a variety of variably hydrothermally metamorphosed mafic ophiolitic rocks from Cyprus, Fe-Mn hydroxyoxide rich chemical sediments (umbers) from Cyprus and Syria¹¹, general fields for mid-ocean ridge basalts¹⁰ general fields for seawater obtained by Pb isotopic analysis of manganese nodules^{12,13}, and sulphides from 21°N on the East Pacific Rise¹⁴. The numerals near the symbols for the ophiolitic mafic rocks refer to the samples listed in Table 1.

compositions of the associated basaltic rocks¹⁴. However, Fig. 2 shows that the lead isotope ratios for particular deposits in Cyprus seem to have restricted compositional ranges. In fact, the differences between ratios for Alestos and Mousoulos, both of which are in Cyprus, are actually greater than the differences between Alestos and the EPR 21°N samples. These sulphide deposit fields presumably, therefore, reflect a combination of the local lead isotope and mixing characteristics of each individual ore-forming geothermal system. The influence of source region is well shown by the more radiogenic Pb in the Kuroko volcanic sulphide ore deposits of Japan¹⁶. $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios for galenas from 10 deposits of 18.42–18.62, 15.55–15.72 and 38.51–39.14 are interpreted to reflect variable lead derivation from Miocene volcanics and Palaeozoic basement¹⁶.

One explanation of the differences in the Pb isotopic compositions of the EPR and Cyprus sulphides relates the facts that the EPR deposits are young and small. Hence, they are relatively enriched in unradiogenic leachable basaltic lead. The Cyprus deposits on the other hand, represent a completed process with a higher integrated water: rock ratio. The basaltic lead leached initially may have been diluted by more radiogenic seawater lead added subsequently.

We conclude that: (1) The Pb isotopic composition of the ophiolitic mafic rocks of the Troodos Complex is comparable with that of mid-ocean ridge basalts, but extends to such an unradiogenic combination of $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios as 16.620, 15.401 and 36.511 (T74). (2) The hypothesis of metal derivation by leaching of ophiolitic mafic rocks to form the volcanogenic sulphide ore deposits of Cyprus has some support. (3) The lead is, in fact, a mixture and seems to contain a variable component derived from seawater as well as a basaltic component. This observation differs from observations on recent sulphides from the East Pacific Rise, 21°N (refs 14,15). (4) Pb mixing probably occurred during convection within oceanic crust and not during discharge, as most of the samples examined were deliberately obtained from stockworks which were originally located at variable depths below the ocean floor. Data for a metadolerite sample containing originally dissolved Pb in a hydrothermal epidote-sulphide vein (CY/75/71) are consistent with this suggestion. (5) Lead in conformable sulphide deposits which has been used to construct lead isotope evolution curves (see refs 17,18) need not be exclusively of mantle derivation even when the associated igneous rocks are dominantly basaltic in composition as in the Troodos ophiolitic complex. Contamination by seawater may account for some of the scatter observed in detailed examinations of such conformable deposits (see ref. 19). The Pb isotope scatter is small, however, compared with the overall Pb isotope evolutionary range.

E.T.C.S. acknowledges receipt of operating grant A6114 from the Natural Sciences and Engineering Research Council of Canada and N. H. G. acknowledges partial support of this research by the NERC and by the SRC. We thank Roy Goodwin, Martin Humm (analysts), Sublash Shanblag (draftsman) and Brian O'Donovan (photographer) for invaluable assistance.

Received 2 October 1981; accepted 11 January 1982.

- Desmet, A. *et al.* *Nature* **273**, 527–530 (1978).
- Hekinian, R., Fevrier, M., Bischoff, J. L., Picot, P. & Shanks, W. C. *Science* **207**, 1433–1444 (1980).
- Spiess, F. N. *et al.* *Science* **207**, 1421–1433 (1980).
- Spooner, E. T. C. *Geol. Ass. Can. Spec. Pap.* **20**, 685–704 (1980).
- Spooner, E. T. C. & Fyfe, W. S. *Contr. Miner. Petrol.* **42**, 287–304 (1973).
- Parmentier, E. M. & Spooner, E. T. C. *Earth planet. Sci. Lett.* **40**, 33–44 (1978).
- Spooner, E. T. C. & Bray, C. J. *Nature* **266**, 808–812 (1977).
- Spooner, E. T. C., Chapman, H. J. & Smewing, J. D. *Geochim. cosmochim. Acta* **41**, 873–890 (1977).
- Doe, B. R. & Zartman, R. E. in *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (ed. Barnes, H. L.) 22–70 (Wiley, New York, 1979).
- Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63–87 (1978).
- Gale, N. H., Spooner, E. T. C. & Potts, P. J. *Can. J. Earth Sci.* **18**, 1290–1302 (1981).
- Chow, T. J. & Patterson, C. C. *Geochim. cosmochim. Acta* **26**, 263–308 (1962).
- Reynolds, P. H. & Dasch, E. J. *J. geophys. Res.* **76**, 5124–5129 (1971).
- Vidal, Ph. & Clauer, N. *Earth planet. Sci. Lett.* **55**, 237–246 (1981).
- Brevart, O., Dupré, B. & Allegre, C. J. *Econ. Geol.* **76**, 1205–1210 (1981).
- Fehn, U. *Geol. Soc. Am. A. Meet. Abstr. with Prog.* **11**, 424 (1979).
- Stanton, R. L. & Russell, R. D. *Econ. Geol.* **54**, 588–607 (1959).
- Cumming, G. L. & Richards, J. R. *Earth planet. Sci. Lett.* **28**, 155–171 (1975).
- Koeppl, V. *U.S. geol. Surv. Open-File Rep.* **78**–101, 223–226 (1978).
- Pankhurst, R. J. & O'Nions, R. K. *Chem. Geol.* **12**, 127–136 (1973).
- O'Nions, R. K. & Pankhurst, R. J. *Earth planet. Sci. Lett.* **21**, 13–21 (1973).
- Arden, J. & Gale, N. H. *Analyt. Chem.* **46**, 2–8 (1974).

Temporal changes in the inorganic carbon system of the south-eastern Bering Sea during spring 1980

L. A. Codispoti*, G. E. Friederich*, R. L. Iverson† & D. W. Hood‡

* Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine 04575, USA

† Florida State University, Tallahassee, Florida 32306, USA

‡ PO Box 57, Friday Harbor, Washington 98250, USA

Previous studies of the inorganic carbon system in the southeastern Bering Sea¹ demonstrated the occurrence of extremely low partial pressures of carbon dioxide (p_{CO_2}) and depressed total carbon dioxide concentrations (ΣCO_2) during late spring. To test the hypothesis that these conditions develop during the spring phytoplankton bloom, and to provide biological production data that would be independent of the carbon-14 technique², we monitored the inorganic system in this region

intensively during spring 1980. We demonstrate here that there is a clear relationship between the changes in the inorganic carbon system and the spring phytoplankton bloom. Initially (March–early April), p_{CO_2} values in the surface waters were near or above the mean atmospheric content of 339 p.p.m., and ΣCO_2 concentrations were $\sim 2 \text{ mmol l}^{-1}$. By the end of the bloom, these variables were sometimes below 125 p.p.m. and 1.75 mmol l^{-1} respectively, and the changes were correlated with nutrient decreases and dissolved oxygen and chlorophyll increases. At a mid-shelf station, it was possible to calculate net community organic carbon production from our data. The average rate between 23 April and 23 May was $3.6 \text{ g C m}^{-2} \text{ day}^{-1}$, in fair agreement with carbon-14-based primary production estimates.

Our p_{CO_2} data were taken continuously by sampling an air stream with an intake located near the *T. G. Thompson's* bow and a water stream taken from a depth of 3 m. The water stream was passed through an equilibrator³, and both the air and equilibrator gases were dried before analysis with a Beckman 315B IR analyser. Standards consisted of mixtures of CO_2 and artificial air. Water stream results were corrected for a temperature rise ($\sim 0.5^\circ\text{C}$) between the sea chest and the equilibrator and for an effect that arises when dissolved oxygen partial pressures are different than atmospheric. The p_{CO_2} data shown in Fig. 2 are for moist air. We determined ΣCO_2 ($\text{CO}_2 + \text{H}_2\text{CO}_3 + \text{HCO}_3^- + \text{CO}_3^{2-}$) by stripping acidified samples with nitrogen and drying the evolved gas before determination with a Beckman 215B IR analyser standardized with sodium carbonate solutions. Carbon-14 primary productivity rates were obtained from 8 h incubations in a deck incubator about 1.5°C warmer than ambient surface water. Daily versus incubation period light observations were used to extrapolate daily rates from the incubation period results. The rest of our methods are described elsewhere⁴.

Figure 1 shows the PROBES (processes and resources of the Bering Sea shelf) line and the location of three quasi-stationary fronts that are normal features^{4,5,6}. Although tidal flows are substantial, mean flows in the PROBES region are only $\sim 1\text{--}5 \text{ cm s}^{-1}$. There is evidence⁵ to suggest that the mean flow in the region between the middle and inner fronts is 1 cm s^{-1} , making this a favourable region for inferring biological rates from our p_{CO_2} and ΣCO_2 time series.

Figure 2 shows the temporal evolution of the surface p_{CO_2} regime along the PROBES line in relation to the mean atmospheric value. Initial and final nitrate values are also indicated. These values do not show the presence of the middle front which is confined to the lower half of the water column⁶.

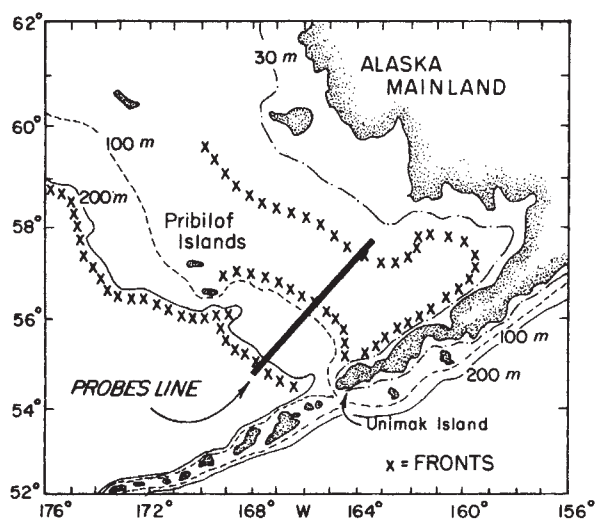


Fig. 1 Location chart showing the approximate position of the fronts that are present in the Bering Sea during the biologically active season, and the position of the main PROBES line.

As atmospheric exchange, vertical mixing, warming of surface waters as spring progressed and precipitation of calcium carbonate would all act to increase p_{CO_2} , the major factor causing the undersaturations seems to be the biological production of organic carbon. This hypothesis is supported by the nitrate data (Fig. 2) and also by the temporal changes in dissolved oxygen and chlorophyll. Initial (late March–early April) dissolved oxygen saturations in the surface waters were near 100%, but they rose as the bloom progressed. On the PROBES line, the maximum saturation value of 141% coincided with the minimum p_{CO_2} value of 120 p.p.m. (position 15 in Fig. 2, 16–17 March). Sometimes, maximum oxygen saturations occurred before minimum p_{CO_2} values. These situations might be expected because of the higher rate of O_2 exchange across the sea surface⁷.

Initial chlorophyll values were generally less than $1 \mu\text{g l}^{-1}$ and never more than $1.5 \mu\text{g l}^{-1}$, but during the bloom, values in excess of $30 \mu\text{g l}^{-1}$ were encountered. Maximum chlorophyll values also occurred before minimum p_{CO_2} values in some cases. This is not surprising as the net productivity rate required to exceed our estimated atmospheric carbon dioxide invasion rates ($\sim 0.7 \text{ g C m}^{-2} \text{ day}^{-1}$, see below) is low in comparison with the maximum production rates.

Although biological production had the greatest effect on the p_{CO_2} values, physical processes also had observable effects. For example, the temporary reversal of the decreasing p_{CO_2} trend at positions 11 and 12 on 22–23 April occurred during a period of enhanced vertical mixing.

As the p_{CO_2} decreases (Fig. 2) seem to arise mainly from biological fixation of carbon, they should provide some insight into the timing of the spring bloom across the PROBES line. They suggest that some net productivity had occurred in the inner frontal region before our first observations. However, the first large p_{CO_2} undersaturations occurred in or near the middle front and the bloom appears to have ended here and at position 15 first. By the end of our observational period (early June), the bloom appeared to be over everywhere, except for portions of the outer shelf and adjoining basin.

Some idea of the temporal variability in ΣCO_2 is given in Fig. 3. These changes, when combined with other data, allow the net community production rates to be estimated. We selected PROBES line position 12 for doing this because it is located in the region of weak mean currents mentioned above, and because we can compare the results with the relatively abundant carbon-14-based primary production measurements at this location. Salinity gradients in this region were weak, and using values that were normalized for salinity would not have changed the results. As the chemical results of biological production can be distributed by vertical mixing and advection, ΣCO_2 values were integrated over the total water column depth.

The interpretation of the changes in ΣCO_2 at position 12 is complicated by the existence of an anomalous station with high subsurface ΣCO_2 concentrations on 16 May. Examination of the temperature and nutrient distributions suggests that this was a transient introduced and removed by advection, but it significantly affects estimates of the ΣCO_2 removals arising from biological production. Regressions of $\Sigma\text{CO}_2 \text{ m}^{-2}$ against time were $3.40 \text{ g C m}^{-2} \text{ day}^{-1}$ ($r^2 = 0.81$) with this value included and $4.32 \text{ g C m}^{-2} \text{ day}^{-1}$ ($r^2 = 0.94$) without it for the 23 April–23 May period. We believe that the latter regression is a better estimate of the change in ΣCO_2 concentrations at position 12 during this portion of the spring bloom as it minimizes random errors without being affected by the advective transient. That this transient influenced only a small portion of the shelf is suggested by similar regressions for the region between positions 11 and 13. Regressions with this value included and deleted gave rates of $4.12 \text{ g C m}^{-2} \text{ day}^{-1}$ ($r^2 = 0.97$) and $3.94 \text{ g C m}^{-2} \text{ day}^{-1}$ ($r^2 = 0.97$) respectively.

The above values have to be adjusted for the effects of mixing and advection, but only horizontal processes have to be considered as we are employing vertically integrated values. The existing transport models^{6,8} demand that a transient such as the

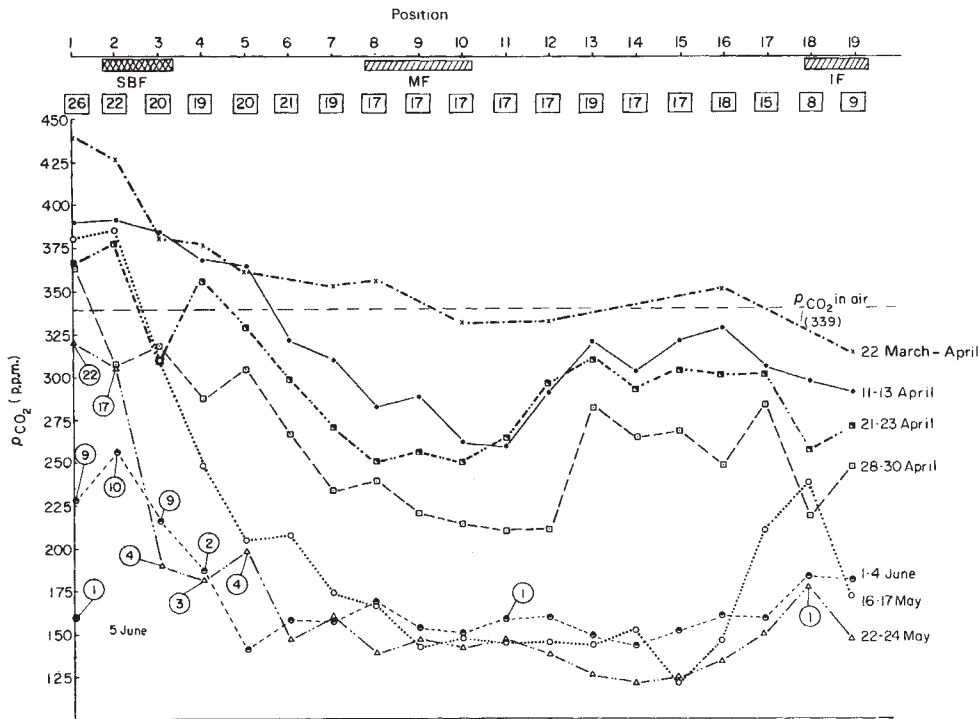


Fig. 2 The temporal evolution of $p\text{CO}_2$ plotted against the mean atmospheric $p\text{CO}_2$ value along the main PROBES line during our 1980 experiment. The approximate locations of the fronts are indicated by the hatched boxes, and initial nitrate concentrations at the surface are indicated in the squares. Surface nitrate concentrations between 22 May and 5 June are shown in circles except for those cases where concentrations were less than $1 \mu\text{g-atom l}^{-1}$. For clarity, partial lines taken on 2-5 and 6-9 May have not been included. Inclusion of these data would not have changed the general trends, but the number of temporary reversals in the generally decreasing trend would have increased. The $p\text{CO}_2$ values are for moist air. On a dry air basis, the mean atmospheric value would be 341 p.p.m.

one observed at position 12 be averaged out over sufficiently long spatial or temporal domains, and it is not possible to do this using only the data for position 12. Hence, we used the position 11-13 space for this calculation (with the transient included). The above models suggest that longshore transport is insignificant, a conclusion that agrees with the longshore ΣCO_2 gradients that we observed. They also parametrically include mixing and advection in a cross-shelf diffusion coefficient which the most recent study suggests is $1 \times 10^6 \text{ cm}^2 \text{ s}^{-1}$ in this domain (L.K. Coachman, personal communication). Applying this value to the appropriate gradients gives a diffusive supply of $0.14 \text{ g C m}^{-2} \text{ day}^{-1}$. Combining this value with the companion result from the regression analysis gives a total rate of $4.3 \text{ g C m}^{-2} \text{ day}^{-1}$. To test the strength of this calculation, we examined every possible combination of regression values and diffusive transports for position 12 only and for the position 11-13 region using calculations with and without the transient. This procedure had little effect on the answer for the position 11-13 region, the maximum difference being $0.3 \text{ g C m}^{-2} \text{ day}^{-1}$. For position 12 only, where it may not be appropriate to calculate the diffusive supply with the available model, the values ranged from 3.21 to $4.86 \text{ g C m}^{-2} \text{ day}^{-1}$.

With the appropriate constants⁹ and our temperature, salinity, $p\text{CO}_2$ and ΣCO_2 data, we calculated that approximately one-third of this decrease was due to the precipitation of calcium carbonate. This value seems reasonable⁷, and if it applies to the entire water column, the decrease arising from the production of organic matter is $2.9 \text{ g C m}^{-2} \text{ day}^{-1}$ (based on our best estimate of $4.3 \text{ g C m}^{-2} \text{ day}^{-1}$ for the total decrease).

A correction also has to be made for CO_2 invasion from the atmosphere. This was done by using the measured $p\text{CO}_2$ gradients across the air-sea interface and the 'piston velocity' concept^{7,10}. Literature values¹¹ were used to convert $p\text{CO}_2$ values into concentrations, and the average wind speed observed during the interval spanned by each gradient observation was used to obtain the appropriate 'piston velocity'. The average time-weighted invasion rate was $0.7 \text{ g C m}^{-2} \text{ day}^{-1}$ for both position 12 and the position 11-13 region. The net community organic carbon production rate arising from the above analysis is $3.6 \text{ g C m}^{-2} \text{ day}^{-1}$.

Seven carbon-14-based primary production rate determinations made at position 12 (between 20 April and 26 May) give

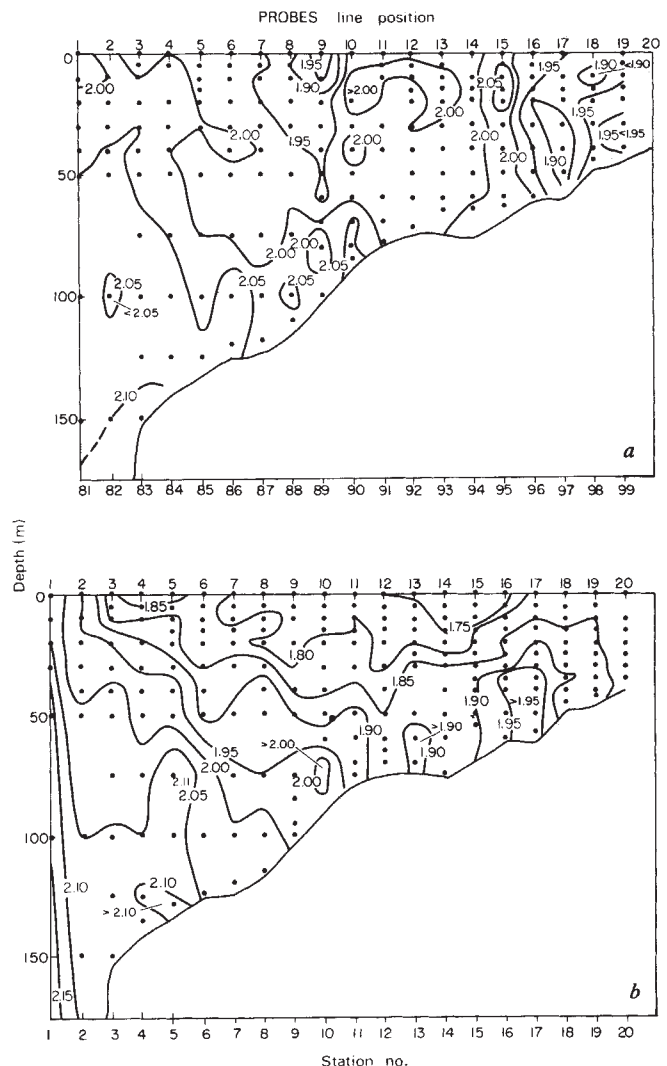


Fig. 3 Total carbon dioxide (in mmol l^{-1}) distribution on the PROBES line, 21-23 April (a) and 22-24 May (b).

a time-weighted average rate of $3.5 \text{ g C m}^{-2} \text{ day}^{-1}$ for the 23 April–23 May period. One would expect that the rate based on the changes in the inorganic carbon system should be smaller than the carbon-14-based rate. The difference between primary and community production could be small in the prevailing conditions, however, because the phytoplankton community at position 12 was composed of diatom species which did not seem to be significantly affected by herbivores that were present in low numbers. Although the carbon-14 method does not always compare favourably with other methods^{12–14}, it does give reasonable estimates of particulate organic carbon production for some diatom populations¹² in eutrophic conditions^{15,16}. This discrepancy could also arise from errors in our comparison. Large errors in the choice of a diffusion coefficient, in the air–sea gas exchange, or in the amount of calcium carbonate precipitation would be required to alter our net community production rate by $1 \text{ g C m}^{-2} \text{ day}^{-1}$, but our test of our calculation suggests that factors such as insufficient resolution of temporal and

spatial gradients could cause errors of this magnitude. An examination of the carbon-14 results suggests that this value may also be a partial function of sampling density. Because differences of the order of $1 \text{ g C m}^{-2} \text{ day}^{-1}$ could arise from errors associated with sampling density, there is no firm evidence to suggest that the carbon-14 method itself suffered from large errors in the conditions of our experiments. However, it is possible that this method estimated net community production more accurately than it estimated primary production.

We thank our colleagues in the PROBES program for data and comments, and G. Grunseich, A. Hafferty and D. Lowman for assistance. Financial support was provided by the NSF's Division of Polar Programs. We particularly thank Dr J. J. Kelly for advice and some necessary equipment. Complete listing of our data can be obtained from PROBES Data Management, c/o Dr H. J. Niebauer, IMS, University of Alaska, Fairbanks, Alaska 99701, USA.

Received 28 August 1981; accepted 11 January 1982.

- Hood, D. W. in *The Eastern Bering Sea Shelf: Oceanography and Resources* Vol. 1 (eds Hood, D. W. & Calder, J. A.) 347–358 (US Department of Commerce, Washington DC, 1980).
- Weichert, G. *Meteor. Forsch.-Ergebnisse*, **22**, 79–86 (1980).
- Schemel, L. E. & Didini, L. A. *U.S. Geol. Surv. Open-File Rep.* 79–273 (1979).
- Iverson, R. L. et al. in *Ecological Processes in Coastal and Marine Systems* (ed. Livingston, R. J.) 437–468 (Plenum, New York, 1979).
- Kinder, T. H. & Schumacher, J. D. in *The Eastern Bering Sea Shelf: Oceanography and Resources* Vol. 1 (eds Hood, D. W. & Calder, J. A.) 53–76 (US Department of Commerce, Washington DC, 1980).
- Coachman, L. K. et al. in *Stratified Flows, Second IAHR Symposium* Vol. 2 (Tapir, Trondheim, in the press).
- Broecker, W. S. *Chemical Oceanography* (Harcourt Brace Jovanovich, New York, 1974).
- Coachman, L. K. & Walsh, J. J. *Deep-Sea Res.* **28A**, 819–846 (1981).
- Millero, F. J. *Geochim. cosmochim. Acta* **43**, 1651–1661 (1979).
- Peng, T. H. & Broecker, W. *Limnol. Oceanogr.* **25**, 789–796 (1980).
- Skirrow, G. in *Chemical Oceanography* Vol. 2 (eds Riley, J. P. & Skirrow, G.) 1–181 (Academic, London, 1975).
- Li, W. K. W. & Goldman, J. C. *Microbiol. Ecol.* **7**, 113–121 (1981).
- Peterson, B. J. *Limnol. Oceanogr.* **23**, 179–183 (1978).
- Peterson, B. J. *A. Rev. ecol. Syst.* **11**, 359–385 (1980).
- Eppley, R. W. & Sloan, P. R. *J. Fish. Res. Bd Can.* **22**, 1083–1097 (1965).
- Ryther, J. W. & Menzel, D. W. *Limnol. Oceanogr.* **10**, 490–492 (1965).

Variability in the abundance of animal and plant species

R. M. Anderson, D. M. Gordon*, M. J. Crawley & M. P. Hassell

Department of Pure and Applied Biology, Imperial College, London University, London SW7 2BB, UK

* Department of Zoology, King's College, London University, London WC2R 2LS, UK

A striking and consistent observation in ecology is that variability in population abundance of a species and average population density are related in both space and time^{1,2}. Taylor and his colleagues have shown that this relationship between variance (V) and mean (M) conforms well to a simple power law for a very wide range of animal and plant species, with the logarithms of sample variance and average sample density following a linear relationship^{1–3} ($V = \hat{a}M^{\hat{b}}$ where $\hat{a}$ and $\hat{b}$ are constants). They suggest that the slope of this line, $\hat{b}$, which generally falls within the range 1 to 2, is a species-specific characteristic reflecting the balance between the opposing behavioural tendencies to aggregate within, and migrate from, centres of population density⁴. Taylor and co-workers^{1–6} have stressed that migration between patches is rarely a random process, on the evidence of the observed 'density-dependent' relationships between the degree of dispersion (the variance to mean ratio, V/M and mean abundance per patch). We show here that the relationship between variability and average abundance can be a simple, and inevitable, consequence of chance demographic events in the dynamics of population growth and decline. The precise form of these relationships is determined by the relative magnitude of the various rate processes which govern the dynamics of population change (the birth, death, immigration and emigration rates), and by the degree of spatial and temporal heterogeneity. We do not need to invoke any complex behavioural mechanisms to explain observed patterns.

Random or chance events play a major role in the population dynamics of all organisms. In any given time interval there is

only a certain probability that an organism will die or give birth, or that an immigrant will arrive or emigrant leave, partly as a consequence of the intrinsically discrete structure of populations⁷. The significance of such demographic stochasticity^{7–9} to observed patterns of variability and abundance can be examined using simple markovian population models. We begin by considering the temporal dynamics of a population in a defined space subject to constant rates of birth, death, emigration and immigration (denoted respectively by λ , μ , γ and δ) which are independent of population density (λ , μ and γ are defined as per capita rates while δ is defined per population). A continuous time stochastic model predicts that the probability distribution of population size at any point in time is overdispersed and negative binomial in form^{10–12}, where at time t (the initial population size N_0 at time $t = 0$ is taken to be zero), the mean $M(t)$, and variance $V(t)$, are given by

$$M(t) = \frac{\delta}{r} [e^{rt} - 1] \quad (1)$$

$$V(t) = M(t) \left[1 + \frac{\lambda}{\delta} M(t) \right] \quad (2)$$

The intrinsic growth rate of the population, r , is defined as $r = (\lambda - \mu - \gamma)$. The variance to mean ratio, $V(t)/M(t)$, is greater than unity and tends to rise linearly as average abundance increases (density dependent in the terminology of Taylor^{1,4}), with slope λ/δ . At low population densities, when the value of δ is large in relation to λ (often the case during the early stages of habitat colonization), the variance will be approximately equal to the mean. As average density rises, the log-log plot of variance against mean will become linear, with a maximum slope of 2 (Fig. 1a). The slope of the best-fit linear equation will therefore lie between 1 and 2 depending on the range of densities sampled and the relative magnitudes of the parameters λ and δ (Fig. 1a). Note that the vast majority of observed relationships also have slopes lying between 1 and 2 (Fig. 2).

Simple stochastic models of pure population processes show how different patterns of dispersion can be generated. Birth rates, for example, generate over-dispersion, immigration rates generate randomness, while death and emigration rates tend to create under-dispersion^{8,13,14}. Observed patterns of

dispersion therefore depend on the dynamic balance between these opposing forces.

Stochastic models can also aid in the interpretation of spatial patterns of variability. For example, assume that a habitat is divided into P patches where the immigration rate per patch, δ , is random but dependent on the net rate of emigration from all other patches. When all other rates are constant and indepen-

dent of population density, the mean abundance per patch, $M(t)$, and the variance in density between patches, $V(t)$, are now approximately given by

$$M(t) = N_0 \exp[(A - B)t] \quad (3)$$

$$V(t) = \frac{(A + B)}{(A - B)} M(t) \left[\frac{M(t)}{N_0} - 1 \right] \quad (4)$$

The parameters A and B are defined for notational convenience as $A = \delta P \gamma / (b + \delta P) + \lambda$ and $B = (\mu + \gamma)$; N_0 is the total population size in the habitat at time $t = 0$ and b is the per capita death rate of the migrating organisms. It is assumed in the derivation of equations (3) and (4) that there is a risk associated with migration such that $b \gg \mu$ and that in any given patch population growth is a simple birth (at a rate A)–death (at a rate B) process where a component $(\delta P \gamma / (b + \delta P))$ of the birth rate is the arrival of immigrants from the remaining patches^{8,10}. The probability distribution of abundance per patch is overdispersed in form (similar, but not identical, to the negative binomial distribution) where V/M increases linearly with mean abundance, with a slope of $(A + B)/[N_0(A - B)]$. The log-log plot of variance against mean is again approximately linear, with a slope in general lying between 1 and 2, depending on the range of densities sampled and the values of the population parameters (Fig. 1b). (At very low average densities the slope may even be greater than 2.)

This result does not support Taylor's view (see above) that such patterns are likely to be due to nonrandom migration because our simple model assumes migration to be a random process between patches. Similar conclusions have also been reached by Hanski¹⁵.

We have so far assumed that the rates of population change are independent of density; of course, this is rarely the case in natural habitats. Density dependence, whether acting on the birth, death, emigration or immigration rates (or any combination of these), serves to regulate population abundance in a given habitat to a typical equilibrium K (refs 16–18). Many factors may account for such density dependence, including complex behavioural processes (which often influence emigration and/or immigration rates). When a population is at or near K , density-dependent processes (which reduce birth and immigration rates, or increase death and emigration rates, as density rises) tend to generate underdispersion, as illustrated by the Monte Carlo simulation experiments recorded in Fig. 1c, d. The pattern of dispersion changes from random when the habitat is being colonized, to overdispersed as the population grows exponentially, to random or underdispersed as the population approaches K . A curvilinear relationship between $\log V$ and $\log M$ is thus produced, as shown in Fig. 1d.

Patterns of this form are seldom recorded even for 'K-strategist'¹⁹ species living in relatively stable habitats. One example is provided by the vole, *Microtus pennsylvanicus*, in its grassland habitat, where the degree of dispersion changes from overdispersed to underdispersed as population density increases²⁰. In the case of 'r-strategists'¹⁹ we are unaware of any examples similar to that in Fig. 1d. Thus, contrary to the predictions of our model, there is little evidence that the degree of dispersion declines at high population densities^{1–4}.

The models examined so far, however, fail to take account of a crucial feature of almost all natural habitats, namely, environmental heterogeneity. Because of fluctuating environmental conditions, both in space and time, population rate parameters are in reality random variables, not constants^{21,22}. Environmental stochasticity⁹ acts to compound the underlying variability created by demographic events, and has a central role in determining observed patterns of population dispersion. To illustrate this, we consider a model in which the carrying capacity, K , of each patch is a random variable (normally distributed with mean $\bar{K}$ and variance s_K^2). Monte Carlo simulation studies reveal that log-log plots of the variance in abundance between patches against the mean population density will tend to be approximately linear over all densities,

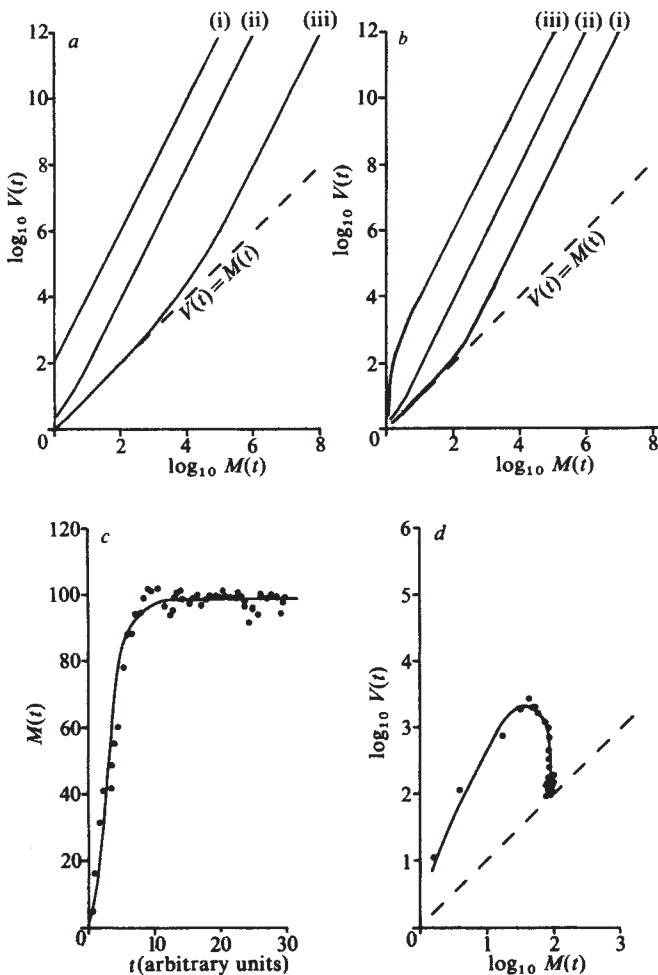


Fig. 1 a, The relationship between the logarithms of the variance in population size at time t , $V(t)$, and mean population size, $M(t)$, predicted by equations (1) and (2) in the text. The three solid lines denote the relationship for different birth (λ) and immigration (δ) rates: line (i), $\lambda/\delta = 0.001$; line (ii), $\lambda/\delta = 1$; line (iii), $\lambda/\delta = 100$. The probability distribution of population size is negative binomial in form¹², with dispersion parameter k (which varies inversely with the degree of dispersion) equal to δ/λ . The dashed line denotes the relationship for a Poisson distribution where $V(t) = M(t)$. b, Similar to a but representing the relationship predicted by the spatial model defined in equations (3) and (4). The three solid lines denote the relationship for different values of the parameter combination $C = [A + B]/[A - B]$ (see text): line (i), $C = 0.01$; line (ii), $C = 1.0$; line (iii), $C = 100$. Initial population size N_0 was set at 1. The dashed line is as defined for a. c, d Present the results of a Monte Carlo simulation experiment of the growth of a population subject to constant rates of birth, immigration and emigration, and a density-dependent rate of mortality $\mu(N)$, where $\mu(N) = \mu + \alpha N$ (N represents population density). The experiments were carried out along standard lines^{13,14} where the equivalent deterministic differential equation for the temporal dynamics of N is of the form $dN/dt = \delta + (\lambda - \gamma)N - (\mu + \alpha N)N$. a Records changes in the mean population size, $M(t)$, through time, estimated from 25 replicate simulations. The solid dots are experimental results while the solid line is fitted by eye to denote the general temporal trend. In b the logarithms of the variances in population size (estimated from the replicate simulations) at time t , $V(t)$, are plotted against the logarithms of means $M(t)$. The dots are experimental results while the solid line is fitted by eye. The dashed line denotes the Poisson prediction where $V(t) = M(t)$. Note that the relationship between $\log_{10} V(t)$ and $\log_{10} M(t)$ is curvilinear, where the degree of dispersion rises to a peak and then declines at high average densities. The parameter values used in the simulations were $\delta = 0.5$, $\lambda = 7.5$, $\mu = 2.0$, $\gamma = 0.5$ and $\alpha = 0.05$ (all defined per time unit).

with a slope lying between 1 and 2, provided the habitat is moderately heterogeneous. There is little evidence of a decline in the degree of dispersion at high average densities, despite the fact that density-dependent processes are regulating overall abundance (Fig. 3a, b, e, f). When sampling is restricted to patches where the density is near to the average carrying capacity, $\bar{K}$, then the slope of the log-log plot of variance against mean ranges from close to, or less than, unity if the habitat is relatively homogeneous, to greater than 2 if the habitat is highly heterogeneous (Fig. 3). Such slopes greater than 2 have occasionally been observed in studies of spatial and temporal trends in the pattern of population dispersion (Fig. 2).

The range of values over which the slope can vary is small (roughly 0.5–2.5, Fig. 2), its precise value being critically dependent on sample size, the range of densities over which samples are collected and the degree of environmental heterogeneity at the time and place of sampling (Fig. 3). For a fixed set of parameter values, for example, the slope may vary greatly between one simulation and another even when sample size is large. The slope will tend, on average, to be determined by the values of the population parameters that control the growth of

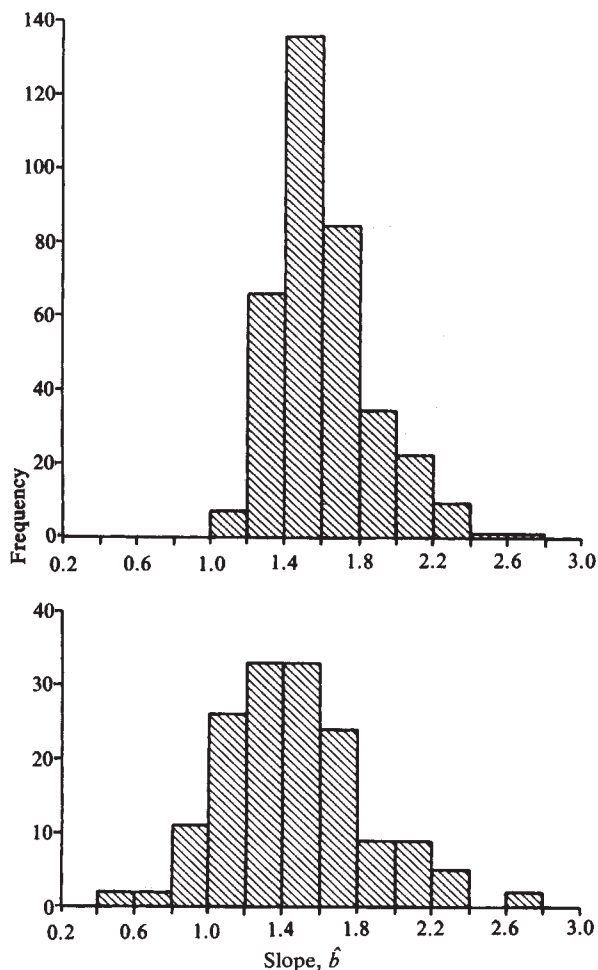


Fig. 2 Frequency distributions of the slopes of the best-fit linear model to the logarithms of sample variance and mean density estimated from a wide range of ecological data. The linear model is of the form $\log_{10} V = \log_{10} \hat{a} + \hat{b} \log_{10} M$ where $\hat{b}$ is the slope and $\log_{10} \hat{a}$ the intercept^{1,4}. *a*, Temporal trends. This figure denotes the frequency distribution of slopes ($\hat{b}$) estimated from samples of the abundances of various moth and aphid species collected over a period of at least 6 yr at sites in Great Britain and mainland Europe (data analysed by Taylor and Woiwood²). The means and variances in population abundance from which the slopes were estimated were calculated from samples taken at one site over a period of 1 yr. The frequency distribution represents slopes for 360 species. *b*, Spatial trends. Similar to *a* but representing the frequency distribution of slopes estimated from samples of the abundances of species ranging from plants, through protozoa, to mammals, taken at one point in time at different spatial sampling sites (data analysed by Taylor, Woiwood and Perry¹). The frequency distribution records slopes for 155 species.

a given species, but these often vary on a regular seasonal basis. It may be possible, however, to discern general trends in the sense that species with high birth rates that live in unstable heterogeneous habitats (*r*-strategists) will tend to have steeper slopes than those that have low birth rates and live in stable

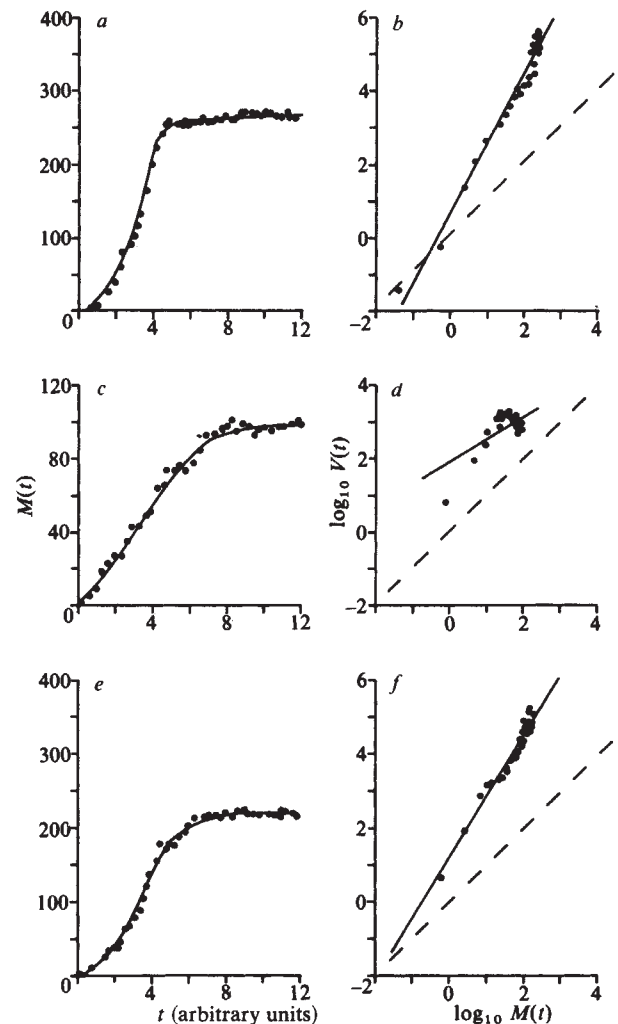


Fig. 3 Monte Carlo simulation studies of the impact of environmental stochasticity on population growth and the degree of dispersion. The results presented represent three experiments using a model in which the carrying capacity of the patches within a habitat is assumed to be a random variable. The model is identical to that described in the Fig. 1 legend to (c, d) except that the density-dependent coefficient, α , was assumed to be normally distributed with mean $\bar{\alpha}$ and variance s_{α}^2 . The simulations were carried out along standard lines^{13,14} except that for each replicate of population growth within a patch a pseudorandom number was used to select a value for α and hence determine the carrying capacity, K , of the patch. K was assumed to be constant through time for any given patch. *a*, *b* Portray the results of one experiment in which the parameter values were set at $\delta = 0.5$, $\lambda = 7.5$, $\mu = 2.0$, $\gamma = 0.4$, $\bar{\alpha} = 0.05$, $s_{\alpha}^2 = 0.025$. *a* Records changes in the mean population density per patch, $M(t)$, through time, of 25 patches. The solid dots denote experimental results while the solid line is fitted by eye. *b* Records the relationship between the logarithms of the variance in population size between patches at time t , $V(t)$ (given that each patch has a different carrying capacity, K , where

$$K = [\lambda - (\gamma + \mu) + [(\gamma + \mu - \lambda)^2 + 4\alpha\delta]^{1/2}] / (2\alpha)$$

and the mean population size per patch, $M(t)$ (sample size = 25). The solid dots denote experimental results while the solid line represents the best-fit linear model of the form $\log_{10} V(t) = \hat{a} + \hat{b} \log_{10} M(t)$, where the constants have the values $\hat{a} = 0.76$, $\hat{b} = 1.91$. The dashed lines denote the Poisson prediction where $V(t) = M(t)$. *c*, *d* Are similar to *a* and *b* except that in the simulation experiments (25 replicates) the variance of the random variable α is set at $s_{\alpha}^2 = 0.0125$ to represent a fairly homogeneous habitat (in the experiments recorded in *a* and *b* the larger value of s_{α}^2 implies a more heterogeneous habitat). In *d* the parameters of the best-fit linear model are $\hat{a} = 1.87$, $\hat{b} = 0.58$. *e*, *f* Are again similar to *a* and *b* except that the parameters λ and μ were changed to $\lambda = 10.0$ and $\mu = 5.0$. The degree of environmental heterogeneity was the same as that for *a* and *b*. In *f* the parameters of the best-fit linear model are $\hat{a} = 1.11$, $\hat{b} = 1.69$.

and relatively homogeneous habitats (*K*-strategists). Some support for this conjecture is provided by the available empirical evidence^{1,4}.

The patterns of dispersion generated by simple models of population growth lead us to believe that observed relationships between variability in population abundance and average density, highlighted by the work of Taylor³, are in general a simple and inevitable consequence of demographic and environmental stochasticity. It is not necessary to invoke explanations based on the behavioural tendencies of species to aggregate and migrate⁴ in order to understand the trends that have been observed in natural habitats. Indeed, in a spatially uniform world, and in the complete absence of demographic stochasticity, such tendencies will not generate power law relationships between population variability (*V*) and average abundance (*M*). When density-dependent factors are of limited significance (*r*-strategists) demographic stochasticity alone is sufficient to account for the approximately linear relationship between the logarithms of variance and mean abundance and for the slopes of such relationships lying on average between 1 and 2 (Fig. 2). In the presence of strong density dependence (*K*-strategists), a degree of environmental heterogeneity (either in space or time, or both) will ensure that such relationships remain approximately linear over all average densities (Fig. 3).

We thank M. Loevinson, J. Soberon, J. Lawton and T. R. E. Southwood for helpful discussions and comments on the manuscript.

Received 16 October 1981; accepted 28 January 1982.

1. Taylor, L. R., Woiwod, I. P. & Perry, J. N. *J. Anim. Ecol.* **47**, 383–406 (1978).
2. Taylor, L. R. & Woiwod, I. P. *J. Anim. Ecol.* **49**, 209–224 (1980).
3. Taylor, L. R. *Nature* **189**, 732–735 (1961).
4. Taylor, L. R. & Taylor, R. A. *J. Nature* **265**, 415–421 (1977).
5. Taylor, R. A. & Taylor, L. R. in *Population Dynamics* (eds Anderson, R. M., Turner, B. D. & Taylor, L. R.) 1–27 (Blackwell, Oxford, 1979).
6. Taylor, L. R., Woiwod, I. P. & Perry, J. N. *J. Anim. Ecol.* **48**, 289–304 (1979).
7. Bartlett, M. S. *Stochastic Population Models in Ecology and Epidemiology* (Methuen, London, 1960).
8. Bailey, N. J. J. *The Elements of Stochastic Processes* (Wiley, New York, 1964).
9. May, R. M. *Stability and Complexity in Model Ecosystems* (Princeton University Press, 1973).
10. Bartlett, M. S. *An Introduction to Stochastic Processes* (Cambridge University Press, 1955).
11. Kendall, D. G. *Ann. Math. Stat.* **19**, 1 (1948).
12. Anscombe, F. J. *Biometrika* **37**, 358–382 (1950).
13. Pielou, E. C. *An Introduction to Mathematical Ecology* (Wiley, New York, 1969).
14. Cox, D. R. & Miller, H. D. *The Theory of Stochastic Processes* (Methuen, London, 1965).
15. Hanski, I. *Oikos* **34**, 293–310 (1980).
16. Krebs, C. J. *Ecology: The Experimental Analysis of Distribution and Abundance* 2nd edn (Harper & Row, New York, 1978).
17. Lack, D. *The Natural Regulation of Animal Numbers* (Oxford University Press, 1954).
18. Hassell, M. P., Lawton, J. H. & May, R. M. *J. Anim. Ecol.* **45**, 471–486 (1976).
19. Southwood, T. R. E. in *Theoretical Ecology, Principles and Applications* (ed. May, R. M.) 26–48 (Blackwell, Oxford, 1976).
20. Grant, P. R. & Morris, R. D. *Can. J. Zool.* **49**, 1043–1052 (1971).
21. Levins, R. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1061–1065 (1969).
22. Lewontin, R. C. & Cohen, D. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1056–1060 (1969).

A possible case of hypervitaminosis A in *Homo erectus*

A. Walker*, M. R. Zimmerman† & R. E. F. Leakey‡

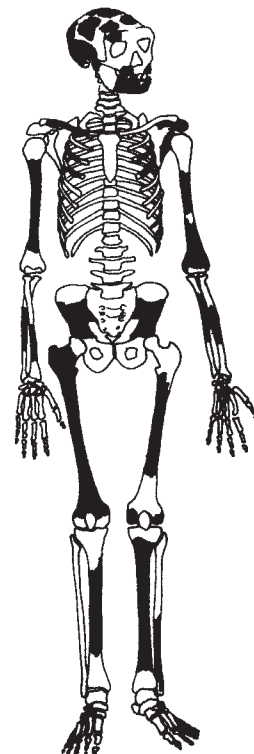
* Department of Cell Biology and Anatomy, The Johns Hopkins University, School of Medicine, Baltimore, Maryland 21205, USA

† Department of Pathology and Laboratory Medicine, Hahnemann Medical College and Hospital of Philadelphia, Philadelphia, Pennsylvania 19102, USA

‡ National Museums of Kenya, Box 40658, Nairobi, Kenya

Following an initial discovery by Bw. K. Kimeu in 1973, sieving operations have recovered the most complete *Homo erectus* skeleton so far known (Fig. 1) from the Upper Member of the Koobi Fora Formation in Area 103, Koobi Fora, East Lake Turkana in Kenya^{1,2}. The partial skeleton shows pathological changes consistent with chronic hypervitaminosis A. We attribute this disorder to the high dietary intake of animal liver, most probably that of carnivores, during a period when the dietary habits of *Homo erectus* were changing.

Fig. 1 Diagram showing the preserved (shaded) parts of *Homo erectus* specimen KNM-ER 1808.



The geological age of the specimen (KNM-ER 1808) is constrained by the ages of the overlying Koobi Fora Tuff and the top of the underlying Olduvai Event. The Koobi Fora Tuff is actually a complex body of tuff, tuffaceous sediment and sediment, as are the Okote and Middle Ileret Tuffs. Cerling and Brown (personal communication) have established that some parts of each of these tuff complexes can be correlated. Fitch and Miller³ reported K/Ar dates on sanidine separates from the Koobi Fora Tuff ranging from 0.53 to 4.44 Myr, dates on the Okote tuff ranging from 0.87 to 1.70 Myr, and a date of 1.48 Myr for the Middle Ileret Tuff. Their preferred ages for the Koobi Fora and Okote Tuffs are 1.57 and 1.56 Myr respectively. Curtis (personal communication) has dated the Okote Tuff at 1.48, 1.46 and 1.44 Myr. Thus a reasonable younger limit on the age of the specimen is 1.5 Myr. The age of the top of the Olduvai Event is given as 1.76 Myr by MacDougall⁴, providing an older limit for the age of the specimen.

The precise position of the polarity transition is uncertain, but is known to lie at least 12 m below the specimen, and may lie much lower in the section. Since the specimen lies fairly close to the Koobi Fora Tuff, a reasonable estimate of its age is 1.6 ± 0.1 Myr. This adult individual is presumed female by comparison of innominate parts with KNM-ER 3228 (ref. 5) and Olduvai Hominid 28 (ref. 6), which are considered male and female respectively by modern human osteological criteria.

The appendicular skeleton shows striking pathology, consisting of subperiosteal diaphyseal deposit of coarse-woven bone. The new bone, 7.0 mm thick in places, thins towards the metaphyses. There is minimal endocranial involvement. Ground thin sections of the tibial shaft (Fig. 2) show pathology confined to the outermost cortex, which has given rise locally to new bone. The sharply demarcated, coarse-woven new bone contains enlarged, sub-spherical and randomly placed lacunae. There is no evidence of abnormal remodelling of the underlying bone. The dense mineralization of the specimen accounts for the fine histological preservation, but precludes satisfactory radiological examination.

Although a disease that no longer exists or has changed its manifestations can neither be diagnosed nor excluded, we suggest that KNM-ER 1808 had chronic hypervitaminosis A. Although the night blindness of vitamin A deficiency has been known since antiquity, the deleterious effects of excessive inges-

tion became known only much more recently. Early polar explorers reported the development of an acute toxic state on ingestion of polar bear, seal or husky dog liver^{7,8}. However, it was not until 1942 that Rodahl and Moore⁹ identified vitamin A as the toxic component of seal and polar bear liver.

Acute hypervitaminosis A, characterized by symptoms of peeling of the skin, vomiting and diarrhoea, headache and convulsions, may be fatal. The condition, now rarely seen in adults, is fairly common in infants after accidental ingestion or misguided parental enthusiasm¹⁰.

Chronic hypervitaminosis A is a considerably more subtle disease, which became recognized as a clinical entity only in the mid-twentieth century. Although the condition is fairly common in children¹¹⁻¹⁴, only 17 cases have been reported in adults since 1951^{7,11,15-22}.

Skeletal manifestations have been variable in these patients. Periosteal calcification of the long bones, well known in children²³, was detected in only 3 of the 17 adults^{7,17,18}. The only case biopsied⁷ showed thickened tibial periosteum with the formation of poorly calcified coarse-fibred new bone. All cases of hypervitaminosis A are probably associated with bone changes, although standard radiographic techniques are insensitive to early changes²⁴. Another case¹⁹, biopsied in the absence of radiological abnormalities, showed large osteocytic lacunae, focal hypermineralization and an increased rate of bone turnover. The enlarged lacunae are thought to be due to the osteocytes assuming an osteoclast-like function, accounting for the hypercalcaemia of hypervitaminosis A.

Bone changes are seen in adult experimental animals with administration of excess amounts of vitamin A^{25,26}. Rats fed large doses of vitamin A showed considerable periosteal bone deposition, with low density, cancellous-type new bone overlying the cortical bone²⁷. Large osteocytes, seen within confluent lacunae, matured rapidly, tending to produce osteolysis with partial breakdown of the bone matrix.

Other conditions causing abnormal calcification must be considered in a differential diagnosis. Hypervitaminosis D leads to

hypercalcaemia and metastatic visceral calcification, with the bones becoming markedly osteoporotic. Cystic lesions of the bones ('brown tumours') are seen in hypoparathyroidism. Some modern cases of hypoparathyroidism have shown increased cortical density, but it is unlikely that a hominid living 1.6 Myr ago would have survived the associated tetany long enough to have developed skeletal changes. Fluorosis also produces osteosclerosis, but this is said to be due to both periosteal and endosteal bone formation. The histology is not well documented^{28,29}. Subperiosteal calcification in scurvy, usually metaphyseal rather than diaphyseal, is associated with cortical thinning and epiphyseal separation²⁸. Infantile cortical hyperostosis involves the mandible and resorbs with ageing³⁰. Generalized cortical hyperostosis (Van Buchem's disease) affects many parts of the skeleton, but shows endosteal thickening, especially in the skull³⁰. The periosteal bone deposition of osteomyelitis variolosa, associated with childhood smallpox, is most prominent near the joint capsule, with possible sequestration of the epiphysis³¹⁻³³. In syphilitic osteomyelitis, the skull is almost always involved; endosteal proliferation is usual, as are large destructive foci (gummas)^{30,31}. Hypertrophic osteoarthropathy is often associated with periosteal bone deposition, but the histology is that of normal cortical bone³¹.

The most likely diagnosis is, therefore, hypervitaminosis A. If we are correct, how did this *Homo erectus* ingest such large doses of the vitamin?

In the Koobi Fora succession, the first record of the use of stone artefacts and their association with animal bones, presumably representing early hominid meals, antedates KNM-ER 1808 by some 200,000 yr. There was probably a major change in the diet of early humans, with a large increase in meat eating, at that period and it may have taken some time to learn which parts of which carcasses were poisonous. One hundred grammes of modern herbivore liver contain 44,000–50,000 IU of vitamin A, whereas 100 g of carnivore liver contain $1.3\text{--}1.8 \times 10^6$ IU as carnivores derive and store large amounts of preformed vitamin from livers of their prey. The condition described above is

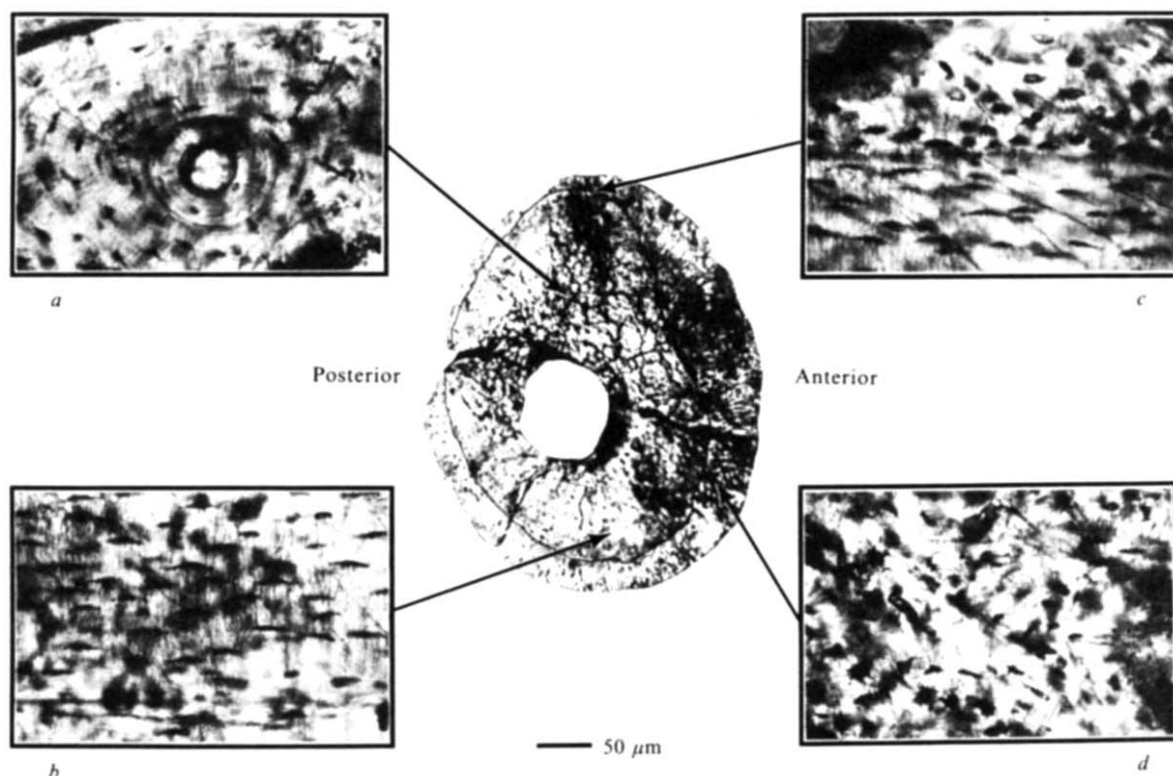


Fig. 2 Cross-section at midshaft of femur showing outer layer of abnormal bone and thin sections taken approximately from the positions shown. *a*, Normal osteon from deeper part of shaft. *b*, Outer lamellar bone showing normal, flattened osteocytic lacunae and canaliculi. *c*, Junction of normal lamellar and abnormal bone, showing sudden transition from normal lacunae to subspherical, disorganized lacunae with more randomly disposed canaliculi. *d*, Disorganized area typical of the abnormal bone.

unlikely to be due to ingestion of the easily masticated liver of a herbivore (there is no evidence of the use of fire at this time), but quite possibly to a diet containing carnivore liver. Interference behaviour of the type seen today between carnivore species³⁴ would be expected when early hominids began to compete directly with carnivores. Some of these interactions might have resulted in carnivores being killed and eaten by hominids.

Alternatively, or additionally, selection for tolerance to high levels of vitamin A may have been slow, with KNM-ER 1808 representing one individual who could not store vitamin A at high levels. This is unlikely, however, because normal human liver levels are within herbivore ranges³⁵. It is more likely that

early humans learned to avoid carnivore liver rather than were selected for new physiological limits.

We thank the Museum Trustees of Kenya, Bw. K. Kimeu, who made the initial discovery and directed the sieving operations, and the following for valuable discussions and help: J. Bieri, J. W. Bowerman, F. H. Brown, T. E. Cerling, J. B. Cohn, G. H. Curtis, M. W. Donner, J. P. Dorst, M. T. Freedman, M. J. Glincher, M. G. Leakey, A. L. Lehninger, O. Lovejoy, L. Perez, S. S. Siegelman and B. Van Valkenburgh. We acknowledge support from the National Geographic Society and the Kenya National Museums to R.E.F.L. and NSF grants BNS 75-16879 and 78-24499 to A.W.

Received 14 July 1981; accepted 25 January 1982.

1. Leakey, M. G. & Leakey, R. E. *Koobi Fora Research Project* Vol. 1 (Clarendon, Oxford, 1978).
2. Leakey, R. E. F. *Nature* **248**, 653-656 (1973).
3. Fitch, F. J. & Miller, J. A. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. & Leakey, R. E. F.) 123-147 (University of Chicago, 1976).
4. MacDougall, I. in *The Earth: its Structure, Origins and Evolution* (ed. McElhinny, M. W.) 543-561 (Academic, London, 1979).
5. Leakey, R. E. F. *Nature* **261**, 574-576 (1976).
6. Day, M. H. *Nature* **232**, 363-381 (1971).
7. Gerber, A., Raab, A. P. & Sobel, A. E. *Am. J. Med.* **16**, 729-745 (1954).
8. Shearman, D. J. C. *Br. med. J.* **1**, 283-285 (1978).
9. Rodahl, K. & Moore, T. *Biochem. J.* **37**, 166-168 (1943).
10. Furman, K. I. *Am. J. clin. Nutr.* **26**, 575-577 (1973).
11. Bergen, S. S. & Roels, O. A. *Am. J. clin. Nutr.* **16**, 265-269 (1965).
12. Mahoney, C. P., Margolis, M. T., Krauss, T. A. & Labbe, R. F. *Pediatrics* **65**, 893-896 (1980).
13. Pease, C. N. *J. Am. med. Ass.* **182**, 980-985 (1962).
14. Ruby, L. K. & Mital, M. A. *J. Bone Jt Surg.* **56**, 1283-1287 (1974).
15. Bifulco, E. *Am. J. clin. Nutr.* **16**, 265-269 (1965).
16. DiBenedetto, R. J. *J. Am. med. Ass.* **201**, 700-702 (1967).

17. Frame, B., Jackson, C. E., Reynolds, W. A. & Umphrey, J. E. *Ann. intern. Med.* **80**, 44-48 (1974).
18. Hofmann, K. J., Milne, F. J. & Schmidt, C. S. *Afr. med. J.* **54**, 579-580 (1978).
19. Jowsey, J. & Riggs, B. L. *J. clin. Endocr.* **28**, 1833-1835 (1968).
20. Katz, C. M. & Tzagournis, M. *Metabolism* **21**, 1171-1176 (1972).
21. Stimson, W. H. *New Engl. J. Med.* **265**, 369-373 (1961).
22. Sulzberger, M. B. & Lazer, M. P. *J. Am. med. Ass.* **146**, 788-793 (1951).
23. Haworth, J. C. *et al. Can. med. Ass. J.* **104**, 521-522 (1971).
24. Shaywitz, G. A., Siegel, N. J. & Pearson, H. A. *J. Am. med. Ass.* **238**, 1749-1750 (1977).
25. Clark, I. & Smith, M. R. *J. Biochem. Chem.* **239**, 1266-1271 (1964).
26. Wolke, R. E., Eaton, H. D., Nielson, S. W. & Helmboldt, C. F. *Am. J. Path.* **98**, 677-685 (1969).
27. Bélanger, L. F. & Clark, I. *Anat. Rec.* **158**, 443-452 (1967).
28. Lichtenstein, L. *Diseases of Bones and Joints* (Mosby, St Louis, 1970).
29. Teitelbaum, S. L. in *Pathology* 7th edn (eds Anderson, W. A. D. & Kissane, J. M.) 1905-1977 (Mosby, St Louis, 1977).
30. Edeiken, J. & Hodes, P. J. *Roentgen Diagnosis of Diseases of Bone* (Williams & Wilkins, Baltimore, 1973).
31. Aegerter, E. & Kirkpatrick, J. A. *Orthopedic Diseases* 3rd edn (Saunders, Philadelphia, 1968).
32. Cockshott, P. & MacGregor, M. Q. *Jl Med.* **27**, 369-387 (1958).
33. Davidson, J. C. & Palmer, P. E. S. *J. Bone Jt Surg.* **45B**, 687-693 (1963).
34. Eaton, R. L. *Carnivore* **2**, 9-16 (1979).
35. Moore, T. *Vitamin A* (Elsevier, Amsterdam, 1957).

β -Endorphin is processed differently in specific regions of rat pituitary and brain

S. Zakarian & D. G. Smyth

Laboratory of Peptide Chemistry, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Recent studies have shown that β -endorphin is found in the pituitary with two related peptides, the C'-fragment (β -endorphin₁₋₂₇)^{1,2} and the des-histidine derivative of the C'-fragment (β -endorphin₁₋₂₆)³, and that each of these peptides (Fig. 1) occurs also in an α ,N-acetyl form^{3,4}. The six peptides are derived from a single precursor by differential proteolytic cleavage and acetylation. The fact that different forms of β -endorphin predominate in different regions of the pituitary^{5,6} implies that there are region-specific differences in processing of the precursor. Because β -endorphin is known to occur widely in the brain^{7,8} as well as in the pituitary we have now investigated the distribution of the six peptides in selected regions of rat brain. We report here that, as in the pituitary, different regions of the brain contain strikingly different proportions of the six β -endorphin-related peptides, apparently as a result of differential processing.

Dissected regions of rat pituitary (10 male Sprague-Dawley rats, 250 g) and of rat brain (30 rats), and also whole brain (5 rats), were extracted in acetone/HCl/water (40:1:5, v/v) and the soluble products fractionated by gel filtration on a Sephadex G75 column (70×1.5 cm) in 50% acetic acid. The elution position of the group of β -endorphin-related peptides was identified using ¹²⁵I-labelled bovine markers added to each homogenate. The marker peptides, isolated from ox pituitary⁹, were β -endorphin, α ,N-acetyl β -endorphin, the C'-fragment and α ,N-acetyl C'-fragment. The endogenous β -endorphin-related peptides, together with the markers, were resolved by chromatography on a sulphopropyl Sephadex C25 (pyridinium

form) column (70×0.6 cm) in 50% acetic acid with a 0-1 M pyridine linear gradient (100 ml mixer) and located by radioimmunoassay using an antibody raised against β -endorphin¹⁰. The identity of four of the six peptides was finally confirmed by HPLC on a C₁₈ microbondapak column in 0.01 M hydrochloric acid with a 25-40% acetonitrile linear gradient (30 min)⁵. The elution positions of the immunoreactive peptides were compared with the positions of the corresponding bovine peptides in consecutive chromatograms. Each experiment was repeated several times (five times for regional distribution of the pituitary and two to six times for region of brain); only minor differences were observed in the patterns of distribution and the amounts of each peptide.

In rat pars intermedia six peptides having β -endorphin immunoreactivity were resolved (Fig. 2); four corresponded to the known marker peptides, β -endorphin₁₋₃₁ (VI), acetyl β -endorphin₁₋₃₁ (V), β -endorphin₁₋₂₇ (IV) and acetyl β -endorphin₁₋₂₇ (II), and the remaining components (III and I), which were the principal peptides of the pars intermedia, had chromatographic properties corresponding to porcine β -endorphin₁₋₂₆ and its α ,N-acetyl derivative. Thus in rat pars intermedia, β -endorphin is a minor component and the major β -endorphin-related peptides are essentially inert as opiates. In contrast, the major peptide in the anterior pituitary with the approximate molecular size of β -endorphin eluted in the position of β -endorphin₁₋₃₁. Similar processing patterns have

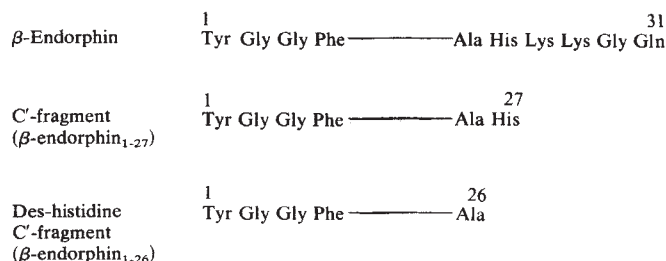


Fig. 1 β -Endorphin-related peptides identified in porcine pituitary. In the rat, valine replaces alanine at position 26 (ref. 11).

been found in other species, including pig^{3,5}, mouse, cat and guinea pig (S.Z. and D.G.S., unpublished results).

The six peptides identified in rat pituitary were shown to be present also in rat brain. Ion-exchange chromatography of the β -endorphin-related peptides extracted from whole brain (Fig. 3) indicated that the acetylated forms were relatively minor; the principal peptides were β -endorphin₁₋₃₁ and β -endorphin₁₋₂₆. Examination of the peptides extracted from regions of brain, however, revealed two contrasting processing patterns, one for the hypothalamus, midbrain and amygdala, and the other for the hippocampus, dorsal colliculae and brain stem. The hypothalamus contained predominantly the biologically potent form of β -endorphin and the midbrain and amygdala again contained β -endorphin but it was accompanied by β -endorphin₁₋₂₆ (Fig. 4a); in these regions there were negligible acetyl peptides. In contrast the hippocampus, dorsal colliculae and the brain stem also contained the α ,N-acetyl forms of

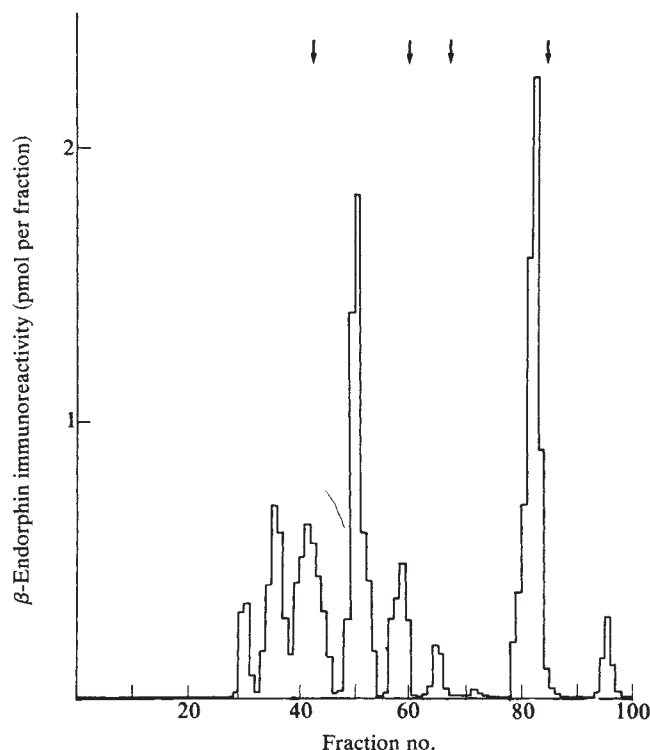


Fig. 3 Ion-exchange chromatography of β -endorphin-related peptides extracted from whole rat brain. The mean weight of intact brain was 1.83 ± 0.17 g (20 rats). The elution positions of the ^{125}I -labelled reference peptides (α ,N-acetyl β -endorphin₁₋₂₇, β -endorphin₁₋₂₇, α ,N-acetyl β -endorphin₁₋₃₁ and β -endorphin₁₋₃₁) are indicated from left to right by the arrows. α ,N-acetyl β -endorphin₁₋₂₆ and β -endorphin₁₋₂₆ do not elute in positions indicated by the marker peptides. Note that the major peptides are β -endorphin₁₋₃₁ and β -endorphin₁₋₂₆. Fractions were 1.5 ml.

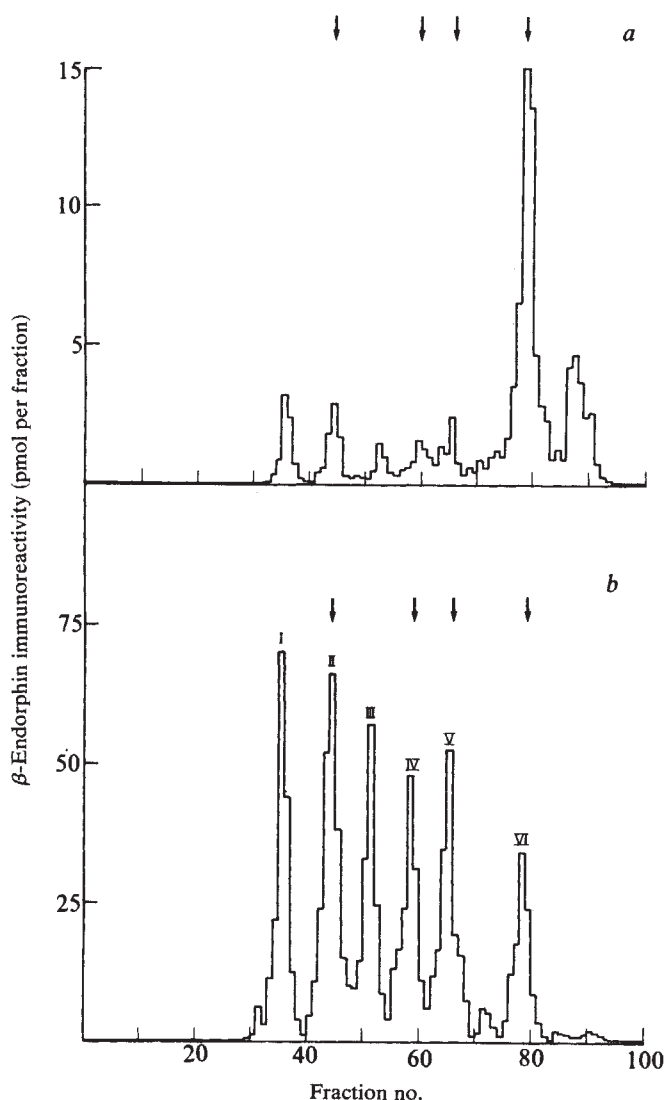


Fig. 2 Ion-exchange chromatography of β -endorphin-related peptides extracted from the anterior pituitary (a) and pars intermedia plus posterior pituitary (b) of the rat. The weights (in mg) of freshly dissected tissues were: anterior pituitary, 8.9 ± 1.2 ; pars intermedia, 1.1 ± 0.1 (mean $\pm$ s.d. calculated from 20 rats). The elution positions of the radio-iodinated reference peptides (α ,N-acetyl β -endorphin₁₋₂₇, β -endorphin₁₋₂₇, α ,N-acetyl β -endorphin₁₋₃₁ and β -endorphin₁₋₃₁), which exhibit slightly greater retention than the endogenous peptides, are indicated respectively from left to right by the arrows. α ,N-acetyl β -endorphin₁₋₂₆ (peak I) and β -endorphin₁₋₂₆ (Peak III) chromatographed in positions not indicated by the marker peptides. The column conditions are described in the text. Fractions were 1.5 ml.

β -endorphin₁₋₂₇ and β -endorphin₁₋₂₆; the NH_2 forms of the peptides were less significant (Fig. 4b).

These results point to the existence of differential processing mechanisms in various regions of rat brain, as is the case in different regions of the pituitary. In the anterior pituitary and hypothalamus, β -endorphin seems to be formed by the action of an enzyme that cleaves between lysylarginine and tyrosine (positions 39–41 of rat lipotropin¹¹). In the amygdala and midbrain, however, β -endorphin₁₋₂₆ is an additional component, which would seem to require the action of two further enzymes, an endopeptidase for cleavage at Lys-Lys-Gly (positions 68–70 of rat lipotropin) and a carboxypeptidase B-like enzyme for removal of C-terminal lysine and histidine. On the other hand, in the pars intermedia, hippocampus, dorsal colliculae and brain stem the most abundant peptides are derivatives of either β -endorphin₁₋₂₇ or β -endorphin₁₋₂₆ and they are largely present in the acetylated form. The elaboration of these peptides would seem to involve the action of at least four enzymes—the two endopeptidases, the carboxypeptidase B-like exopeptidase and an α ,N-acetylating enzyme. Note that of the six peptides related to β -endorphin, β -endorphin alone possesses potent analgesic properties^{12,13} and the formation of any one of the other peptides is accompanied by almost complete loss of the analgesic activity^{3,14}.

The secondary processing of β -endorphin may involve reactions that are time-dependent. In the brain, the processing could occur in the cell bodies where β -endorphin seems to be synthesized, during transport along axons, or during storage at the nerve terminals. Thus the peptides present in the different regions might reflect the duration of their exposure to specific processing enzymes. Indeed, the predominance of β -endorphin₁₋₂₆ and β -endorphin₁₋₂₇ rather than β -endorphin₁₋₃₁ in the brain stem, dorsal colliculae and hippocampus would be

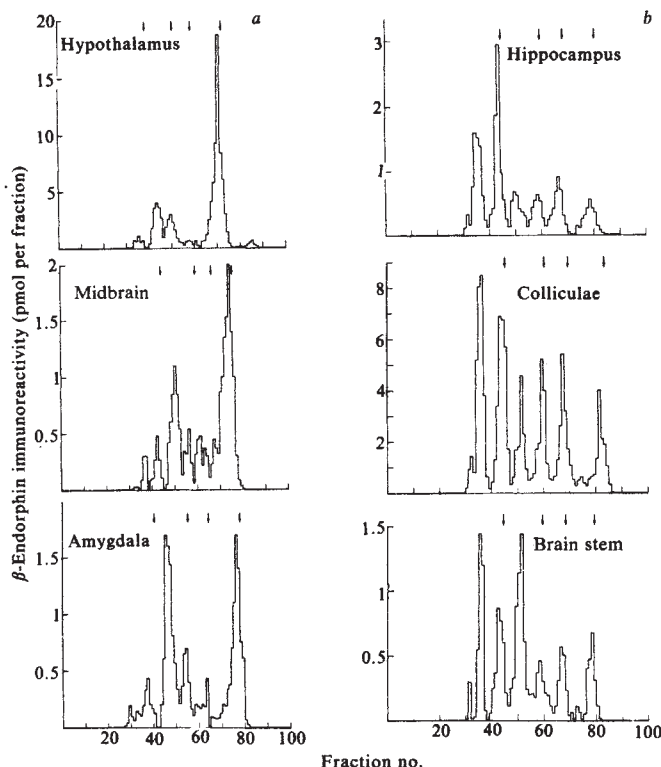


Fig. 4 Ion-exchange chromatography of β -endorphin-related peptides extracted from dissected regions of rat brain. The mean weights (in mg) of the freshly dissected tissues were: hypothalamus, 102 ± 18 ; midbrain, 131 ± 15 ; amygdala, 60 ± 19 ; hippocampus, 153 ± 22 ; dorsal colliculae, 127 ± 11 ; brain stem, 270 ± 30 (values calculated from 30 rats). The elution positions of the ^{125}I -labelled reference peptides (α, N -acetyl β -endorphin $_{1-27}$, β -endorphin $_{1-27}$, α, N -acetyl β -endorphin $_{1-31}$ and β -endorphin $_{1-31}$) are indicated from left to right by the arrows. Fractions were 1.6 ml. Note that the hypothalamus, midbrain and amygdala exhibit one pattern of processing (a) while the hippocampus, dorsal colliculae and brain stem exhibit another (b).

consistent with the concept that most of the secondary processing takes place distally from the hypothalamus, as it is only in the hypothalamus that cell bodies occur^{10,15}. On balance, however, the specific nature of the products formed in the different tissues, and the multiplicity of the enzymes necessary for their formation, suggests that distance from the site of biosynthesis is unlikely to be responsible for the organized distribution of peptides observed. The nature and amounts of the peptides present in each region are probably determined by the controlled action of specific enzymes located in those regions.

We thank Kamela Maruthainar for technical assistance.

Received 7 September 1981; accepted 18 January 1982.

- Bradbury, A. F., Smyth, D. G. & Snell, C. R. in *Peptides, Chemistry, Structure and Biology* (eds Walter, R. & Meienhöfer, J.) 609–615 (Ann Arbor Science, Michigan, 1975).
- Smyth, D. G., Snell, C. R. & Massey, D. E. *Biochem. J.* **175**, 262–270 (1978).
- Smyth, D. G., Smith, C. C. & Zakarian, S. in *Proc. int. Narcotics Conf.*, Kyoto (ed. Takagi, H.) (Kodansha-Elsevier, Tokyo, in the press).
- Smyth, D. G., Massey, D. E., Zakarian, S. & Finnie, M. D. A. *Nature* **278**, 252–254 (1979).
- Smyth, D. G. & Zakarian, S. *Nature* **288**, 613–615 (1980).
- Smyth, D. G., Zakarian, S., Deakin, J. F. W. & Massey, D. E. in *Peptides of the Pars Intermedia* (eds Lawrenson, G. & Evered, D. C.) 79–96 (Pitman Medical, 1981).
- Bradbury, A. F., Feldberg, W. S., Smyth, D. G. & Snell, C. R. in *Opiates and Endogenous Opiate Peptides* (ed. Kosterlitz, H. W.) 9–18 (Elsevier, Amsterdam, 1976).
- Rossier, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 5162–5165 (1977).
- Massey, D. E. & Smyth, D. G. *Biochem. Soc. Trans.* **8**, 751–753 (1980).
- Zakarian, S. & Smyth, D. G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5972–5976 (1979).
- Drouin, J. & Goodman, H. M. *Nature* **288**, 610–612 (1980).
- Loh, H. H., Tseng, L. F., Wei, E. & Li, C. H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2895–2898 (1976).
- Feldberg, W. S. & Smyth, D. G. *Br. J. Pharmac.* **60**, 445–454 (1977).
- Deakin, J. F. W., Doströfsky, J. O. & Smyth, D. G. *Biochem. J.* **189**, 501–506 (1980).
- Bloom, F. E., Battenberg, E., Rossier, J., Ling, N. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1591–1596 (1978).

Insulin-like growth factor I stimulates growth in hypophysectomized rats

Eugen Schoenle*, Jürgen Zapf*, René E. Humbel† & E. Rudolf Froesch*

* Metabolic Unit, Department of Medicine, University Hospital, CH-8091 Zurich, Switzerland

† Biochemical Institute, University of Zurich, CH-8006 Zurich, Switzerland

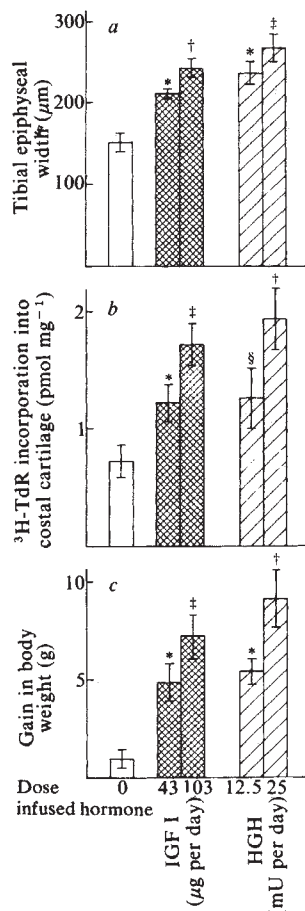
Although growth hormone stimulates the growth of hypophysectomized rats, it has long been proposed^{1,2} that the effects are not direct but instead are mediated by the somatomedin peptides. Two of these are insulin-like growth factor I (IGF I) and insulin-like growth factor II (IGF II)³, so called because they are closely related to insulin in structure^{4,5}. IGF I and IGF II have somatomedin activity *in vitro*^{6,7} but until now insufficient amounts of the peptides have been available to test their *in vivo* activity. We now report that pure IGF I stimulates the growth of hypophysectomized rats in a dose-dependent manner. This strongly supports the notion that the action of growth hormone is mediated by peptides of the somatomedin family.

The *in vitro* effects of insulin-like growth factors have been extensively studied. Both IGF I and IGF II are potent stimulators of cell replication, and of DNA and RNA synthesis in fibroblasts⁷, chondrocytes³ and calvarian cells (ref. 8 and C. Schmid and E.R.F., in preparation). Moreover, they mimic the effects of insulin on insulin target tissues such as adipose tissue⁷ and muscle⁹.

The recent development of specific antibodies against IGF I (ref. 6) and somatomedin C (ref. 10) have made it possible to investigate in man the dependence of these factors on the growth hormone (GH) status in various pathophysiological situations. Available evidence indicates that immunoreactive IGF I and somatomedin C are identical^{11–13}. The following clinical findings support the somatomedin concept²: immunoreactive IGF I is decreased in hypopituitarism and increased in acromegaly⁶. Treatment of hypopituitary children with growth hormone increases the serum immunoreactive IGF I levels¹⁴. Similarly, treatment of hypophysectomized rats with growth hormone causes a rise in their decreased serum IGF level¹⁵. Although clinical findings and *in vitro* results support the notion that IGF rather than growth hormone itself is the true growth stimulating hormone, there is no unequivocal direct evidence for this. There have been recent reports^{16,17} that plasma peptide fractions containing somatomedin activity caused an increase in body weight and sulphate-incorporating activity of cartilage in Snell dwarf mice when administered subcutaneously three times per day for 2–4 weeks. Whereas van Buul-Offers *et al.*¹⁶ also found an increase in nose-to-tail length, Holder *et al.*¹⁷ observed no such effect. To prevent hypoglycaemia in the animals after bolus injections of the somatomedin preparation, van Buul-Offers *et al.*¹⁶ had to add 10% glucose to the drinking water. The latter problem was circumvented by Ellis *et al.*¹⁸, who demonstrated that continuous subcutaneous (s.c.) infusion of a partially purified (4% purity) IGF preparation stimulated growth of hypophysectomized rats. The latter, in contrast to Snell dwarf mice, do not grow under thyroxine treatment, and therefore appear to represent a more specific experimental model for testing the somatomedin hypothesis. Therefore in this study we used hypophysectomized rats.

Male Tif RAI rats (100–120 g) were used for all experiments. Hypophysectomy was carried out 3–5 weeks before starting the experiments. Only those rats whose body weight did not increase by >0.2 g per day were considered to be totally hypophysectomized. The rats had free access to food (Altromin; Lage, FRG) and drinking water, and were kept on a 12-h light and dark cycle. Two doses of pure IGF I (43 and 103 μg per day)

Fig. 1 Stimulation of various growth indices by administration *in vivo* of IGF I and HGH to hypophysectomized rats (*a* = tibial epiphyseal width; *b* = thymidine incorporation into rib cartilage; *c* = daily body weight gain). Hormones were administered to groups of four rats for 6 days by continuous infusion from s.c. implanted minipumps. IGF I was dissolved in 0.1 M acetic acid (43 or 103 μg per day); HGH (12.4 mU or 25 mU per day) (Nanormon (2 U mg^{-1}); Nordisk Insulin Laboratory, Denmark) was dissolved in saline. No difference was found between control rats treated for the same period with 0.1 M acetic acid or saline or left untreated. $\square$, Control hypophysectomized rats; $\blacksquare$, IGF I-treated hypophysectomized rats; ▨ , HGH-treated hypophysectomized rats. Each bar represents the mean $\pm$ s.e.d. ($n=4$). * $P<0.001$; $\dagger P<0.01$ compared with control (Student's *t*-test). $\ddagger P<0.01$; $\S P<0.05$ compared with the lower dose of hormone (Student's *t*-test).



higher dose (103 μg per day) to 168% compared with controls. The effect of IGF I was similar to that obtained in hypophysectomized rats treated with human growth hormone (HGH); whether administered by continuous s.c. infusion or by two intraperitoneal injections daily, HGH (12.5 or 25 mU per day) caused stimulation of growth to 157 and 176%, respectively, of controls.

Figure 1b shows DNA synthesis *in vitro* in costal cartilages of the same rats. IGF I (43 μg per day) increased ³H-TdR incorporation to 171%, and 103 μg per day to 239% compared with controls. Again the effect of HGH was similar: an increase to 172% (12.5 mU per day) and 265% (25 mU per day) was observed. The gain in body weight (Fig. 1c) during IGF I infusion was also comparable to that obtained during HGH administration.

Thus, IGF I administered to hypophysectomized rats mimics the effects of GH on these important indices of growth. Longitudinal growth is stimulated directly by IGF I. Our results strongly support the somatomedin concept²; however, they do not exclude other indirect and direct effects of GH on various tissues.

We thank Mr Meier and Dr Maier for supplying the hypophysectomized rats and Cristina Hauri and Therese Steiner for technical assistance. This work was supported by grant no. 3.380-0.78 from the Swiss NSF

Received 12 November 1981; accepted 2 February 1982.

- Salmon, W. D. & Daughaday, W. H. *J. Lab. clin. Med.* **49**, 825-836 (1957).
- Daughaday, W. H. *et al. Nature* **235**, 107 (1972).
- Froesch, E. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **73**, 2904-2908 (1976).
- Rinderknecht, E. & Humbel, R. E. *J. biol. chem.* **253**, 2769-2776 (1978).
- Rinderknecht, E. & Humbel, R. E. *FEBS Lett.* **89**, 283-286 (1978).
- Zapf, J., Walter, H. & Froesch, E. R. *J. clin. Invest.* **68**, 1321-1330 (1981).
- Zapf, J., Schoenle, E. & Froesch, E. R. *Eur. J. Biochem.* **87**, 285-296 (1978).
- Canalis, E. J. *J. clin. Invest.* **66**, 709-719 (1980).
- Poggi, C., Le Marchand-Brustel, Y., Zapf, J., Froesch, E. R. & Freychet, P. *Endocrinology* **105**, 732-730 (1979).
- Furlanetto, R. W., Underwood, L. E., Van Wyk, J. J. & D'Ercole, A. J. *J. clin. Invest.* **60**, 648-657 (1977).
- Van Wyk, J. J., Svoboda, M. E. & Underwood, L. E. *J. clin. Endocr. Metab.* **50**, 206-208 (1980).
- Hintz, R. L., Liu, F., Mashall, L. B. & Chang, D. J. *J. clin. Endocr. Metab.* **50**, 405-407 (1980).
- Hintz, R. L., Liu, F. & Rinderknecht, E. J. *J. clin. Endocr. Metab.* **51**, 672-673 (1980).
- Kemp, S. F., Rosenfeld, R. G., Liu, F., Gaspich, S. & Hintz, R. L. *J. clin. Endocr. Metab.* **52**, 616-621 (1981).
- Kaufmann, U., Zapf, J. & Froesch, E. R. *Acta endocr.* **87**, 716-727 (1978).
- van Buul-Offers, S. & van den Brande, L. *Acta endocr.* **92**, 242-257 (1979).
- Holder, A. T., Spencer, E. M. & Preece, M. A. *J. Endocr.* **89**, 275-282 (1981).
- Ellis, S., Zapf, J., Froesch, E. R. & Humbel, R. E. *Endocrinology* **108**, Suppl. 103, Abstr. (1981).
- Greenspan, F. S., Li, C. H., Simpson, M. E. & Evans H. M. *Endocrinology* **45**, 455-463 (1949).

³H-noradrenaline release potentiated in a clonal nerve cell line by low-intensity pulsed magnetic fields

Richard Dixey

Department of Medical Electronics, St Bartholomew's Hospital, West Smithfield, London EC1A 7BE, UK

Glen Rein

Department of Chemical Pathology, Queen Charlotte's Maternity Hospital, London W6 0XG, UK

Recent studies have shown that extremely low-frequency (ELF) electric fields¹ and radiofrequency amplitude-modulated electromagnetic fields^{2,3,25} can influence ion fluxes in neural tissue. Furthermore, the application of low-frequency magnetic (LFM) fields in the successful treatment of non-union fractures and pseudo-arthritis has been widely reported⁴⁻⁶. Stimulated by these findings and the wide variety of effects reported in other biological systems⁷⁻¹⁰, we have set out to establish a model system for studying the biological effects of such low-frequency

were administered continuously for 6 days via s.c. implanted Alzet minipumps (no. 2001; Alza). The pumps (mean pumping rate $1.02 \pm 0.15 \mu\text{l h}^{-1}$, mean filling volume $219.2 \pm 8.7 \mu\text{l}$), were implanted under light ether anaesthesia. The pure IGF I preparation used was the same as that isolated and sequenced elsewhere⁴; its specific biological activity, as determined in the rat fat pad assay, was 330 mU mg^{-1} (ref. 7). Pure IGF I was dissolved in 0.1 M acetic acid.

The body weight of the rats was measured daily at 08.00 h and they were killed after 6 days. The two lowest ribs were removed and ³H-thymidine (³H-TdR) incorporation into DNA was determined as follows: segments of cartilage were prepared free of perichondrial tissue and incubated (24 h, 37°C) in Krebs-Ringer-phosphate buffer containing human serum albumin (2 mg ml^{-1}), nonessential amino acids (1% v/v; 5832 Difco), amino acids HeLa (1% v/v; 5790 Difco), glycine (0.2 mM), glutamine (2 mM) and ³H-TdR (1 μM). After incubation the cartilages were dried and weighed, digested in concentrated (90%) formic acid (30 min, 90°C), then counted for radioactivity in 10 ml Instagel (Packard) in a β counter. The tibia test was carried out according to Greenspan *et al.*¹⁹. The concentration of IGF I was determined by radioimmunoassay (RIA) before the minipumps were filled and after they were removed. The IGF I concentration was $1.88 \mu\text{g} \mu\text{l}^{-1}$ in one group of rats before and $1.90 \mu\text{g} \mu\text{l}^{-1}$ after 6 days of infusion. Thus, the IGF I concentration did not decrease in the implanted minipumps during 6 days under the skin.

Blood was sampled by aortic puncture and immunoreactive IGF I in pooled serum was determined by RIA. IGF I concentrations were 168 ng ml^{-1} and 286 ng ml^{-1} , respectively, in the two IGF I-infused groups of rats. No immunoreactive IGF I was detectable in control and growth hormone-treated rats. The endogenous rat IGF, which increases after growth hormone treatment¹⁵, does not cross-react in the RIA with human IGF I. The lower dose of IGF I (43 μg per day) stimulated an increase in the tibial epiphyseal width (Fig. 1a) to 139%, the

fields that readily allows analysis of their mechanism of action. Clonal lines in tissue culture are an obvious choice for such a model, first, because the analysis is simplified by the presence of only one cell type and second, because the geometry of the biological sample can be specified to allow precise quantitation of applied field strengths and induced current densities. One such clonal line, designated PC12, is a likely candidate for such an approach as it expresses many properties of differentiated sympathetic neurones¹¹⁻¹⁶, including calcium-dependent release of neurotransmitters^{11,12,14}. We report here that ³H-noradrenaline (NA) release from PC12 cells is stimulated by an inductively coupled 500-Hz LFM field with a magnitude comparable with certain cholinergic stimuli in this system.

PC12 cells are derived from a transplantable rat pheochromocytoma¹⁵; they synthesize, store and release dopamine, noradrenaline and acetylcholine¹¹⁻¹³ and express sodium action potentials¹⁶. In these experiments, PC12 cells were grown in Dulbecco's modified Eagle's medium containing 5 g l⁻¹ glucose, supplemented with 10% calf serum and 5% horse serum in the absence of nerve growth factor.

A magnetic field was generated by passing current through a pair of concentric 400-turn coils, 10.6 cm in diameter, separated by 14.7 cm. Two culture dishes were placed symmetrically between the coils at heights of 4.95 and 9.75 cm from the bottom coil. The changing magnetic field induces circulating eddy currents in the conducting medium in the dishes, with the current density increasing from zero at the centre to a maximum at a distance approximately equal to the coil radius. The range of currents so induced was limited by containing the cells in an annulus. This was achieved by gluing a small inverted Petri dish into the centre of the 60-mm sample dish and plating the cells in the annular channel so formed. This limits the range of induced current densities to within $\pm 20\%$ of the mean¹⁷. The magnetic field was generated between the coils by wiring them in parallel to a pulse generator and driving circuit similar to that used in clinical trials for the treatment of non-union fractures in Newcastle and London (P. O. Byrne and D. Outram, in preparation). The generator gives a train of square pulses with a range of different pulse durations, frequencies and intensities¹⁷.

For these experiments we used a pulse width of 0.6 ms, with 1.4 ms between each pulse, giving an overall frequency of 500 Hz. The current waveform was smoothed due to the large inductance of the coils (Fig. 1). The applied magnetic field will have the same form as the current in the coils and can be calculated from the geometry of the coils and the current^{17,18}. The maximum field applied to the sample was 8.5×10^{-4} Wb m⁻² (8.5 G) and the minimum 1.6×10^{-4} Wb m⁻² (1.6 G). The rate of change of the magnetic field, calculated from the voltage induced in a search coil, varied between +1.8 and -0.9 Wb m⁻² s⁻¹, and the calculated circulating eddy current density varied between +0.025 and -0.013 A m⁻², inducing an electric field ranging from $+3.8 \times 10^{-2}$ to -1.9×10^{-2} V m⁻¹.

³H-NA release was measured using logarithmic-phase PC12 cells by a modification of the procedure of Greene and Rein¹². After 2 h incubation with ³H-NA, cytoplasmic NA stores were

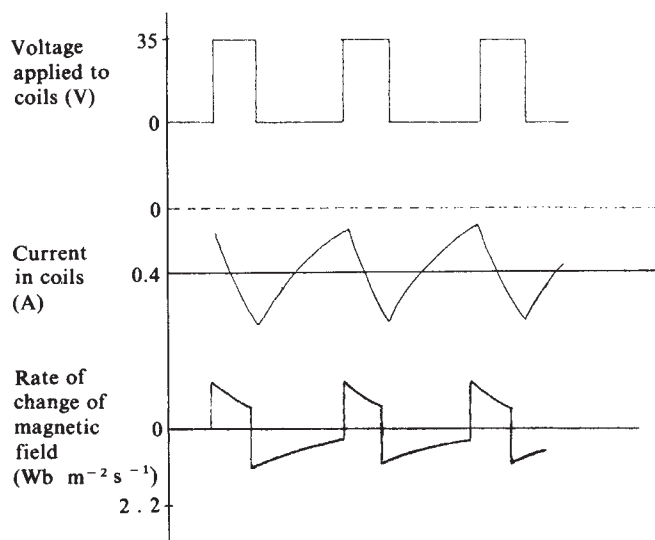


Fig. 1 Waveform generated by the low-frequency magnetic field generator. The magnetic field was calculated from the total measured current I in the coils using the formula

$$B_z = \mu_0 \mu_2 \times \frac{na^2}{2} \left[\frac{1}{(z_1^2 + a^2)^{3/2}} + \frac{1}{(z_2^2 + a^2)^{3/2}} \right]$$

where a = radius of the coils, n = number of turns, and z_1 and z_2 are distances from each coil to point of calculation. $\partial B_z / \partial t$ was calculated from the voltage (V) induced in a search coil of area A and N turns: $V = NA \partial B / \partial t$ (refs 17, 18).

largely depleted as the NA was spontaneously released from the cells during a 3.5-h washout period. Release was then monitored for 12 sequential 15-min periods. At the 105- and 120-min time periods (designated experimental), two dishes were removed from the Numetal shield, which protects the cells from stray magnetic fields, and placed between the coils for a total of 13 min. To eliminate the possibility that LFM fields dislodge cells into the medium, thereby generating artificially high values, medium from the cells was centrifuged at 2,200g for 5 min before assaying for radioactivity. The pellets were solubilized and counted separately.

In all 10 trials in which the LFM fields were applied, release values (c.p.m.) obtained at 105 and 120 min were always higher than the values obtained from the two preceding time periods. To estimate the extent of the observed stimulation, the first six release values for all 22 trials were fitted to the polynomial function $\log Y = \phi(x) + N + T$ from which the control curve of Fig. 2 is obtained; Y represents the c.p.m. in the supernatants, x is the time in minutes, N is a normalizing term allowing direct comparison between dishes and T is a term inserted for the experimental period¹⁹. In the control dishes ($n = 12$) the observed values for the experimental periods fitted within 1 s.d. of those obtained by extrapolating the polynomial function,

Table 1 ³H-noradrenaline release from PC12 cells

c.p.m. expressed as % of cell content	<i>n</i>	Time (min)			
		90	105	120	135
Control	12	0.717 $\pm$ 0.039	0.686 $\pm$ 0.046	0.657 $\pm$ 0.044	0.628 $\pm$ 0.033
Experimental	10	0.717 $\pm$ 0.039	0.875 $\pm$ 0.064	0.838 $\pm$ 0.061	0.628 $\pm$ 0.033
Experimental + 15 mM Mg ²⁺	8	0.717 $\pm$ 0.039	0.673 $\pm$ 0.015	0.686 $\pm$ 0.02	0.628 $\pm$ 0.033
Pellets	10	0.049 $\pm$ 0.012	0.05 $\pm$ 0.014	0.045 $\pm$ 0.013	0.06 $\pm$ 0.024

Release values were generated by regression analysis and expressed graphically as in Fig. 2. In the elevated Mg²⁺ trials the medium for the 105- and 120-min time periods, during field stimulation, contained 15 mM MgSO₄. The aspirated medium was treated as described in Fig. 2 legend. Pellets obtained from spinning aspirated medium (experimental series) at 2,200g for 5 min were solubilized in 1.5 ml of 1 M NaOH and assayed for radioactivity in 10 ml Instagel. All values are expressed as a percentage of the residual intracellular radioactivity and are shown as mean $\pm$ s.e.m.

showing that the washout curve varies smoothly with time. The experimental values (during LFM field stimulation) were elevated above the controls, with the treatment effect being highly significant ($P < 0.001$).

Expressing the stimulated release as the difference between the control and experimental values relative to the control value for a given time period, a percentage stimulation ($\pm$ s.e.m.) of $27.5\% \pm 4.9$ is obtained for both experimental time periods. This elevation is likely to be due to increased release of NA from the cells in monolayer and not to contamination by LFM field-induced floating cells, because the pellet counts for the experimental periods are not significantly different ($P > 0.5$) from those before or after it (see Table 1). The temperature of the coils was continuously monitored and did not rise above the 37°C of the CO_2 incubator.

In eight further trials in which 15 mM Mg^{2+} was included in the medium for the experimental periods, the rate of release in control washout curves did not alter, but the stimulated release was abolished (see Table 1). The stimulated release

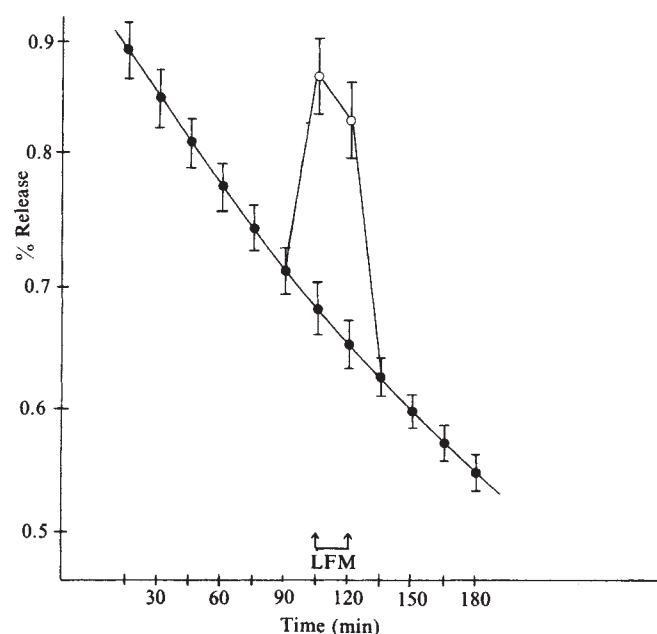


Fig. 2 Effect of a 500-Hz pulsed magnetic field on noradrenaline release from PC12 cells. This curve was generated by regression analysis from a population of 22 curves. Cells were grown in collagen-coated annular dishes made by gluing an inverted 35-mm dish to a 60-mm dish. All manipulations were carried out in a 37°C warm room. Logarithmic cells ($\sim 1 \times 10^6$ cells per dish) were incubated for 2 h in Dulbecco's modified Eagle's medium containing 1 mM (–)sodium ascorbate and $2\text{ }\mu\text{Ci}$ ^3H -(–)NA ($11\text{ Ci}/\text{mmol}$; Amersham) in a CO_2 (7%) incubator maintained at 37°C . The radioactive medium was then removed and the cells washed three times with fresh medium (at 37°C). The cells were incubated in a $37^\circ\text{C}/\text{CO}_2$ atmosphere for an additional 3.5 h in 1.5 ml fresh medium. During this time, the medium was changed after 2 h and then at 30-min intervals. Thereafter the medium (1.5 ml) was changed every 15 min for 12 consecutive periods. The medium from these changes was transferred to polypropylene tubes and centrifuged at $2,200g$ for 5 min to pellet gently any floating cells. The supernatants were then assayed for radioactivity in 10 ml Instagel (Packard). At the end of each experiment, radioactivity remaining in the cells was assayed by scraping the cells sequentially in 0.5 ml 1 M HCl , 0.5 ml 1 M NaOH and 0.5 ml water. The pooled scrapings were then assayed for radioactivity as above. Release values were calculated from the ratio of the c.p.m. in the supernatant, at a given time period, to the total residual intracellular radioactivity at the end of the experiment. The mean total intracellular count ($\pm$ s.e.m.) was $389,018 \pm 26,827$. We have shown previously that 84% of the radioactivity released from the cells is unmetabolized NA¹². ●, Control curve; ○, Stimulated points.

values in the presence of elevated Mg^{2+} were not significantly different from the controls ($P > 0.3$). In three of the trials, following exposure to 51.5 mM K^+ (ref. 11) following the period of stimulation by the LFM, the cells released a substantial proportion of their NA store, indicating that irreversible damage had not occurred.

These results suggest that PC12 cells can be used as a model system for studying the effects of LFM fields on neuronal functions. Release of catecholamines (spontaneous or preloaded with ^3H -NA) from PC12 cells can be induced by depolarizing concentrations of K^+ (refs 11, 14), the sodium ionophore veratridine¹¹, or by activating acetylcholine receptors with nicotine¹² or carbamylcholine¹⁴. Whereas the maximum stimulation of NA release induced by nicotine ($10\text{ }\mu\text{M}$)¹² is $\sim 300\%$ in similar experimental conditions to those used here, the acetylcholine analogue carbamylcholine ($100\text{ }\mu\text{M}$) induced a 60% stimulation of release (L. A. Greene and G.R., unpublished observation). The 27.5% stimulation observed here may not be maximal, as a dose-response curve has not yet been established, but is still of comparable magnitude with that obtained with certain cholinergic stimuli.

Previous studies on Ca^{2+} efflux from chick and cat cerebrum have stressed the importance of a 6–20-Hz frequency 'window'^{20,21}, although an initial report on efflux of γ -aminobutyric acid in the cat cortex showed a small increase induced by a 200-Hz directly coupled signal²². In the experiments reported here release was induced by a 500-Hz pulsatile waveform. Clearly the frequency dependence needs further investigation and studies are under way to determine the effect of different stimulation frequencies and the importance of the pulse shape and modulation.

The observation that Mg^{2+} inhibits the LFM field-induced release from PC12 cells indicates a requirement for extracellular Ca^{2+} , as Mg^{2+} has been shown to be a specific Ca^{2+} antagonist in several biological systems²³. Because depolarization-induced release is also inhibited by Mg^{2+} in PC12 cells^{11,12,14}, it is possible that LFM field-induced release occurs via a similar exocytotic mechanism in which Ca^{2+} fluxes are involved²⁴. However, on biophysical grounds alone, the direct imposition of Ca^{2+} entry by this LFM field is an unlikely hypothesis, because we estimate that the maximum electric field intensity is $\sim 4 \times 10^{-2}\text{ V m}^{-1}$, far less than the 10^3 V m^{-1} field associated with a typical action potential. A more plausible hypothesis is that the cation binding sites on the outer surface of the plasma membrane may be highly sensitive to such weak stimuli²⁰ and that subsequent cooperative changes in Ca^{2+} binding may influence membrane stability, thereby promoting Ca^{2+} entry and vesicular release.

Clearly, this physical stimulus offers a new approach for studying the mechanisms involved in stimulus-secretion coupling²⁴ as well as having important and widespread clinical implications.

We thank Dr Michael Lunt for the design and characterization of the physical approach, Mr Nigel Hathway for assistance in analysing the results and the Smith's Charity for supporting R.D. during this study.

Received 16 September 1981; accepted 26 January 1982.

- Bawin, S. M. & Adey, W. R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1999–2003 (1976).
- Bawin, S. M., Kaczmarek, L. K. & Adey, W. R. *Ann. N.Y. Acad. Sci.* **297**, 74–81 (1975).
- Blackman, C. S., Elder, J. A., Bennane, S. G., Weil, C. M. & Eichinger, D. C. in *Symp. biol. Effects of Electromagnetic Waves*, Airline (1977).
- Lancet* **i**, 815–816 (1981).
- Bassett, C. A., Mitchell, S. N. & Gaston, S. R. *J. Bone Jt Surg.* **A63**, 551–559 (1981).
- Bassett, C. A., Pauluk, R. J. & Pilla, A. A. *Ann. N.Y. Acad. Sci.* **238**, 243–262 (1974).
- Moore, R. L. *Can. J. Microbiol.* **25**, 1145–1151 (1979).
- Tenforde, T. S. (ed.) *Magnetic Field Effects on Biological Systems* (Plenum, New York, 1979).
- Frey, A. H. *Ann. N.Y. Acad. Sci.* **238**, 273–280 (1975).
- 18th Hanford Life Sci. Symp.* Richland, Washington (1978).
- Greene, L. A. & Rein, G. *Brain Res.* **129**, 247–263 (1977).
- Greene, L. A. & Rein, G. *Brain Res.* **138**, 521–528 (1977).
- Greene, L. A. & Rein, G. *Nature* **268**, 349–351 (1977).
- Richie, A. K. *J. Physiol., Lond.* **286**, 541–561 (1979).
- Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2424–2428 (1976).
- Dichter, M. A., Tischler, A. S. & Greene, L. A. *Nature* **268**, 501–504 (1977).

17. Lunt, M. J. & Watson, B. W. *Clin. Phys. physiol. Measurement* (submitted).
18. Lunt, M. J. *Med. Biol. Engng Comput.* (submitted).
19. Armitage, P. *Statistical Methods in Medical Research* (Blackwell, Oxford, 1971).
20. Bawin, S. M., Sheppard, A. & Adey, W. R. *Bioelectrochem. Bioenergetics* **5**, 67–76 (1978).
21. Adey, W. R. *Biosystems* **8**, 163–175 (1977).
22. Kaczmarek, L. K. & Adey, W. R. *Brain Res.* **66**, 537–540 (1974).
23. Simpson, L. L. *J. Pharm. Pharmacol.* **20**, 889–910 (1968).
24. Douglas, W. W. *Br. J. Pharmacol.* **34**, 451–474 (1968).
25. Adey, W. R., *Physiol. Rev.* **61**, 435–514 (1981).

Agonist selectivity and second messenger concentration in Ca^{2+} -mediated secretion

D. E. Knight*, T. J. Hallam & M. C. Scrutton

Departments of Physiology* and Biochemistry, King's College, Strand, London WC2R 2LS, UK

Cells typically carry receptors for several excitatory and/or inhibitory agonists. In many cases the pattern of response observed is characteristic of the particular agonist used^{1–4}. These differences in the nature of the responses cannot always be explained on the basis that the agonists use different intracellular second messengers^{3–5}. For example, in the blood platelet, Ca^{2+} seems to be the second messenger⁶ for such excitatory agonists as ADP, adrenaline, thrombin, collagen and 11-epoxy-methanoprostaglandin H_2 (U-46619). These agonists cause secretion of serotonin, ADP and ATP from the amine storage granules whereas secretion of acid hydrolases from the lysosomes has been reported to occur only on stimulation by higher concentrations of collagen and thrombin^{7–11}. We have investigated this agonist selectivity and report here results suggesting that it is not related to cytosolic Ca^{2+} concentration.

Figure 1a illustrates the dose-response curves observed for release of serotonin, β -N-acetylglucosaminidase and lactate dehydrogenase when washed human platelets are challenged with thrombin. The concentration (EC_{50}) of thrombin required for half-maximal release of serotonin (0.02 U ml^{-1}) is significantly lower than that of β -N-acetylglucosaminidase (0.09 U ml^{-1}), while no significant release of the cytosolic marker enzyme, lactate dehydrogenase, is observed over the range of thrombin concentration tested. In accord with these observations, data from experiments performed in different conditions reveal a nonlinear relationship between the thrombin-induced release of serotonin and β -N-acetylglucosaminidase from intact platelets, with serotonin being preferentially released at low thrombin concentration (Fig. 1b). This selectivity in the response to thrombin could be explained if secretion of acid hydrolases occurs only when the cytosolic Ca^{2+} concentration is higher than is required for secretion from the amine storage granules. We have therefore examined secretion of serotonin, ADP and β -N-acetylglucosaminidase by washed platelets rendered permeable by exposure to an intense electric field, thereby allowing their internal calcium concentration to be varied¹².

Figure 2 shows the release of serotonin, β -N-acetylglucosaminidase and lactate dehydrogenase measured as a function of Ca^{2+} concentration for platelets rendered permeable and then challenged with Ca^{2+} buffers. Release of β -N-acetylglucosaminidase is observed over the same range of Ca^{2+} concentrations as release of serotonin, but in accord with earlier observations¹² there is no significant Ca^{2+} -dependent release of lactate dehydrogenase. The pattern of serotonin and β -N-acetylglucosaminidase release observed in the permeable cells (Fig. 2) contrasts markedly with that observed when the intact cells are challenged with thrombin (Fig. 1a). From the $\log_{10}$ dose-response relationship observed in several experiments using the permeable cells, the EC_{50} for Ca^{2+} -induced release of β -N-acetylglucosaminidase is calculated to be $1.8 \pm$

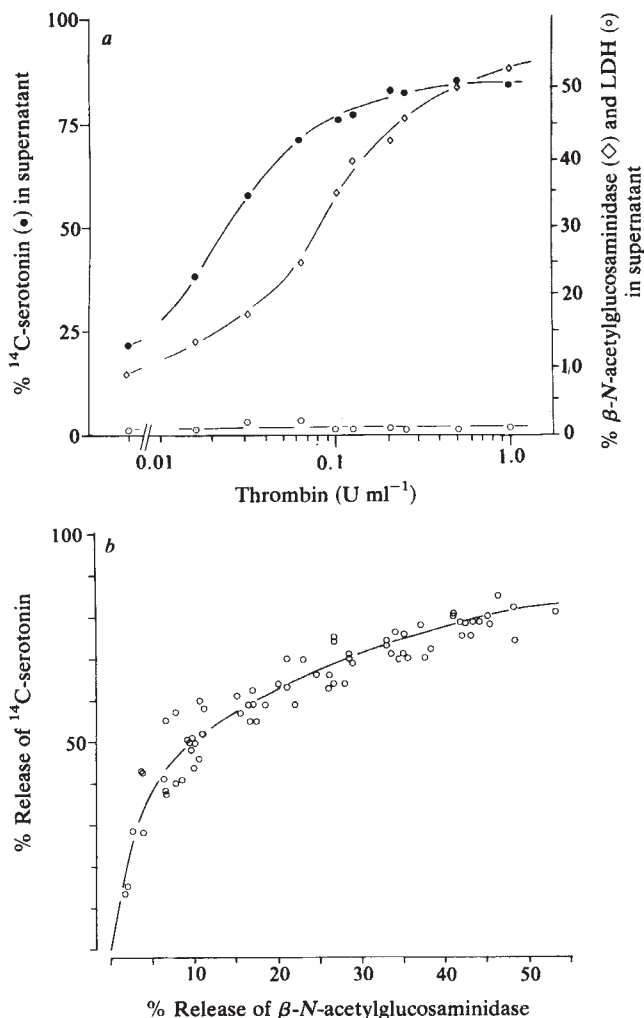


Fig. 1 *a*, The relationship between the extents of release of ¹⁴C-serotonin (●), β -N-acetylglucosaminidase (◇) and lactate dehydrogenase (LDH, ○) and thrombin concentration for intact human platelets. Human blood platelets were suspended in 0.28 M sucrose containing 0.02 M PIPES, 0.01 M K^+ -glutamate, 2 mM MgCl_2 and 0.4 mM EGTA pH 6.6 as described previously¹². Aliquots (0.25 ml) were challenged with various concentrations of thrombin for 5 min at 20 °C before being centrifuged for 2 min at 8,000g. The supernatant fraction was removed and aliquots taken for assay of ¹⁴C content (50 μ l), lactate dehydrogenase (100 μ l) and β -N-acetylglucosaminidase (50 μ l). Lactate dehydrogenase was assayed spectrophotometrically¹⁴ and β -N-acetylglucosaminidase fluorimetrically¹⁵. The ¹⁴C, lactate dehydrogenase and β -N-acetylglucosaminidase contents of the supernatant fractions are expressed as per cent of the total content of the platelet suspension. The total contents of lactate dehydrogenase and β -N-acetylglucosaminidase in the platelet suspension were typically 120 mU ml⁻¹ and 10.0 mU ml⁻¹ respectively. Essentially similar results were obtained when the platelets were suspended in media at pH 7.4 and over a range of Ca-EGTA buffers (10 mM) corresponding to $\sim 10^{-8}$ – 10^{-5} M Ca^{2+} . *b*, The relative amounts of β -N-acetylglucosaminidase compared with ¹⁴C-serotonin released from intact platelets incubated in various concentrations of thrombin and in conditions as in *a*. The amounts released are expressed as a per cent of their cellular content.

0.1 μM ($n = 7$), in good agreement with the value of $1.9 \pm 0.9 \mu\text{M}$ determined previously for serotonin release¹². The EC_{50} is independent of the concentration of EGTA in the Ca^{2+} buffer over the range 5–50 mM, suggesting that this release results from the direct action of micromolar concentrations of Ca^{2+} on the secretory apparatus, rather than through Ca^{2+} -induced release of Ca^{2+} from intracellular stores. The release of β -N-acetylglucosaminidase in this system is smaller than that observed for serotonin, a finding which agrees with the relative

release of β -N-acetylglucosaminidase and serotonin observed both here (Fig. 1) and in previous studies^{7,9} on stimulation of intact platelets by high concentrations of thrombin.

Release of β -N-acetylglucosaminidase is complete within seconds of challenging permeable cells with 10^{-5} M Ca^{2+} , and in a variety of conditions is closely correlated (Fig. 3) with release of both serotonin and ADP, the major nucleotide present in the amine storage granules of human platelets¹³. In contrast, no correlation is observed in these conditions between the release of these constituents and that of lactate dehydrogenase (Fig. 3). In the case of ATP an intermediate situation is observed, as might be expected because significant concentrations of this nucleotide are present in both the cytosol and the amine storage granules of human platelets¹³.

Our data therefore provide further evidence that exocytosis can be induced in human platelets as an immediate consequence of an increase in cytosolic Ca^{2+} concentration and demonstrate

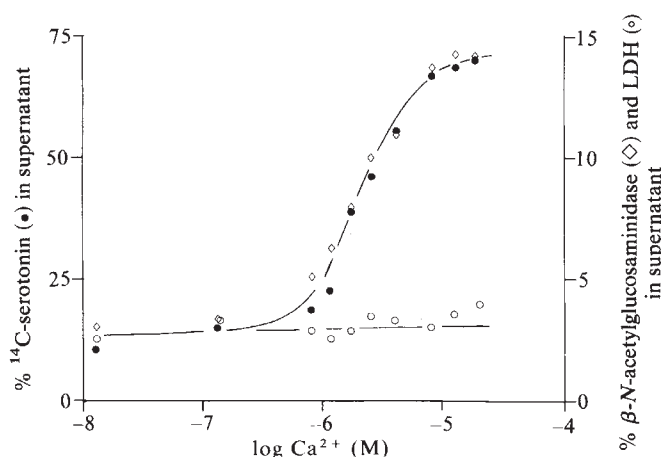


Fig. 2 The relationship between the extents of release of ^{14}C -serotonin (●), β -N-acetylglucosaminidase (◇) and lactate dehydrogenase (○) and the Ca^{2+} concentration in human platelets rendered permeable by exposure to a 20-kV electric field. Human blood platelets were obtained and suspended in 0.28 M sucrose containing 0.2 M PIPES, 0.01 M K^{+} -glutamate, 5 mM ATP, 7 mM MgCl_2 and 0.1 mM acetylsalicylic acid pH 6.6 as described previously¹². After addition of 0.4 mM EGTA the suspension was exposed to an electric field of $20,000 \text{ V cm}^{-1}$ with a time constant of $30 \mu\text{s}$ as described previously¹² and aliquots (0.25 ml) were then immediately challenged with 10 mM Ca-EGTA buffers. The free ionized calcium concentrations were calculated using equilibrium constants of $10^{-5.928}$ M (Ca-EGTA), $10^{-1.445}$ M (Mg-EGTA), $10^{-3.538}$ M (Ca-ATP) and $10^{-3.675}$ M (Mg-ATP)¹⁶. After incubation for 5 min at 22°C the samples were centrifuged for 2 min at 8,000g and the supernatant fractions removed. Aliquots were assayed for ^{14}C -serotonin, lactate dehydrogenase and β -N-acetylglucosaminidase as described in Fig. 1. The contents are expressed as a per cent of the total contents of the platelet suspension. Before exposure to the electric field, 5% of the ^{14}C , and 2% of the lactate dehydrogenase and the β -N-acetylglucosaminidase were present in the extracellular medium. Control studies have shown that: (1) the properties of release of ^{14}C -serotonin and β -N-acetylglucosaminidase do not differ significantly if acetylsalicylic acid is omitted from the suspending buffer; (2) leakage of ATP, which occurs in the presence of 10^{-5} M Ca^{2+} , is complete at the earliest time (30 s) that could be examined after exposure to the electric field; and (3) no significant leakage of ^{14}C -serotonin, lactate dehydrogenase, β -N-acetylglucosaminidase or enolase resulted from exposure to the electric field in the presence of 10^{-9} M Ca^{2+} . These latter observations establish that the membranes of the amine storage granules and the lysosomes are not damaged by exposure to the electric field and that the 'pores' created in the plasma membrane are large enough to allow free passage of molecules of molecular weight <500 but too small to permit leakage of proteins of molecular weight $>80,000$. Further studies are required to define the behaviour in this system of cellular constituents having molecular weights between these two values.

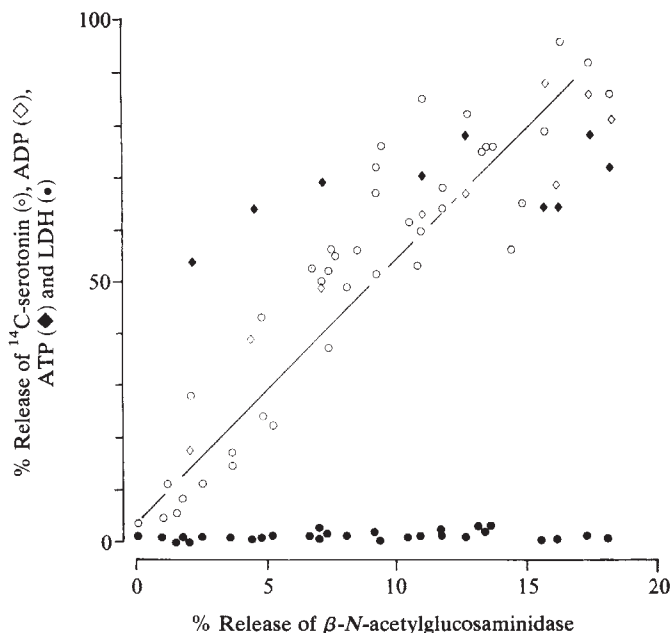


Fig. 3 The relative amounts of β -N-acetylglucosaminidase with ^{14}C -serotonin (○), ADP (◇), ATP (◆) and lactate dehydrogenase (●) released from human platelets incubated in various conditions after exposure to an electric field as described for Fig. 2. Human platelets were prepared, washed and rendered permeable as described for Fig. 2 except that where ADP and ATP were measured, the medium contained 2 mM MgCl_2 and ATP was omitted. ATP and ADP were assayed as described previously¹⁷. The amounts of all components assayed are expressed as per cent of total cellular content. The total contents of ATP and ADP in the platelet suspension were 3.3 and 2.2 μM respectively. After exposure to the electric field in the presence of 10^{-9} M Ca^{2+} the ATP and ADP contents of the supernatant fraction were 2.5 and 0.4 μM respectively. The correlation coefficients between β -N-acetylglucosaminidase and ^{14}C -serotonin, ADP, ATP and lactate dehydrogenase are 0.92 ($n=40$), 0.95 ($n=9$), 0.60 ($n=9$) and 0.37 ($n=33$) respectively. The line shown is a linear regression of ^{14}C -serotonin release on β -N-acetylglucosaminidase release and has a slope of 5.0 ± 0.4 ($\pm \text{s.d.}$).

that the increase in Ca^{2+} concentration required is not significantly different for platelet amine storage granule and lysosomal secretion. Hence, the agonist selectivity observed for secretion of the constituents of these two granules is likely to result from an event in the stimulus-response coupling mechanism other than that responsible for altering the concentration of Ca^{2+} in the platelet cytosol.

Received 7 September 1981; accepted 5 February 1982.

- Petersen, O. H. & Veda, N. J. *Physiol. Lond.* **266**, 819-835 (1977).
- Schramm, M. & Selinger, Z. J. *Cyclic Nucleotide Res.* **1**, 181-192 (1975).
- Putney, J. W. *Pharmac. Rev.* **30**, 209-245 (1978).
- Gardner, J. D. & Jensen, R. T. *Phil. Trans. R. Soc. B* **296**, 17-26 (1981).
- Gardner, J. D. A. *Rev. Physiol.* **41**, 55-66 (1979).
- Feinstein, M. B., Rodan, G. A. & Cutler, L. S. in *Platelets in Biology and Pathology* Vol. 2 (ed. Gordon, J. L.) 437-472 (Elsevier, Amsterdam, 1981).
- Mills, D. C. B., Robb, I. A. & Roberts, G. C. K. *J. Physiol., Lond.* **195**, 715-729 (1968).
- Holmsen, H. & Day, H. J. *Nature* **219**, 760-761 (1968).
- Holmsen, H. *et al. Blood* **46**, 131-142 (1975).
- Kaplan, K. L., Brockman, M. J., Chernoff, A., Leznick, G. R. & Drillings, M. *Blood* **53**, 604-618 (1979).
- MacIntyre, D. E. *Haemostasis* **8**, 274-281 (1978).
- Knight, D. E. & Scrutton, M. C. *Thromb. Res.* **20**, 437-446 (1980).
- Urgurbil, K. & Holmsen, H. in *Platelets in Biology and Pathology* Vol. 2 (ed. Gordon, J. L.) 147-177 (Elsevier, Amsterdam, 1981).
- Bergmeyer, H. U., Bernt, E. & Hess, B. in *Methods of Enzymatic Analysis* (ed. H. U. Bergmeyer) 737-743 (Academic, New York, 1965).
- Dangelmaier, C. A. & Holmsen, H. *Analyt. Biochem.* **104**, 182-191 (1980).
- Martell, A. & Sillen, L. G. *Chem. Soc. spec. Publ.* **17** (1964).
- Hallam, T. J., Scrutton, M. C. & Wallis, R. B. *Thromb. Res.* **21**, 413-424 (1981).

Monoclonal antibody against human IFN- γ

H. K. Hochkeppel

Gesellschaft für Biotechnologische Forschung mbH,
Mascheroder Weg 1, D-3300 Braunschweig, FRG

M. de Ley

Rega Instituut, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Monoclonal antibodies have proved extremely useful for the unambiguous identification, quantification and large-scale purification of a given antigen or a specific epitope. As they are highly specific for a given antigenic determinant, these antibodies are superior to conventional polyclonal antibodies. Since the discovery¹ that mouse spleen B lymphocytes can be fused with mouse B-myeloma cells to yield hybrid cells (hybridomas) secreting specific antibodies derived from the spleen cell fusion partner, several modifications of this method have been used to improve the yield of surviving antibody-secreting hybridomas. We have now included, 5 days after the fusion procedure, an additional step which stabilizes the hybrids against gene loss and improves growth conditions. We report here that, using this method, we have established several stable mouse hybridoma lines secreting monoclonal antibodies to human γ -interferon (HuIFN- γ).

HuIFN- γ was obtained from buffy coats after induction with concanavalin A and subsequently purified by controlled-pore glass adsorption and Sephacryl S200 column chromatography². BALB/c mice were immunized intraperitoneally (i.p.) every third week with 20 units of the HuIFN- γ in a mixture of incomplete Freund's adjuvant and *N*-acetylmuramyl-L-alanyl D-isoglutamine. Three days before fusion the mice were given a final booster (seventh immunization) subcutaneously of 60,000 units of HuIFN- γ in incomplete Freund's adjuvant. The fusion of mouse spleen lymphocytes and mouse (FO) myeloma cells was performed as described by Hochkeppel *et al.*³. After fusion, the hybridomas were first cultured for 5 days in hypoxanthine-aminopterin-thymidine (HAT) medium containing

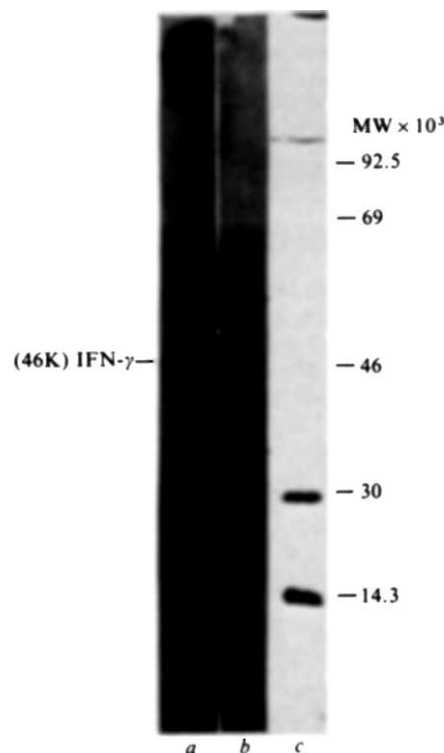


Fig. 2 ELISA using HuIFN- γ and monoclonal anti-HuIFN- γ . Partially purified HuIFN- γ which had been pretreated with 0.1 M SDS alone was separated, in parallel with ¹⁴C-labelled standard proteins, by SDS-polyacrylamide gel electrophoresis according to Laemmli⁵. All proteins were electrophoretically transferred to a nitrocellulose sheet according to Towbin *et al.*⁶. While the section of the blot containing marker proteins was subjected to autoradiography, ELISA was performed on the section containing HuIFN- γ . First the blot was saturated for 2 h with 3% bovine serum albumin (BSA)/5% horse serum to block all the nonspecific binding sites. After washing with Tris-buffered saline, monoclonal anti-HuIFN- γ ($10 \mu\text{g ml}^{-1}$) purified from hybridoma culture supernatants was added at a dilution of 1:50 and incubated at 25°C for 16 h. Subsequently rabbit IgG directed against mouse IgM was added at a 1:200 dilution in 5% horse serum, 0.5% BSA (4 h, 25°C) and finally goat peroxidase-conjugated IgG directed against rabbit IgG (1:500; 2 h at 25°C). The reaction was followed by observing the colour change with *O*-dianisidine used as a substrate for H₂O₂. *a*, HuIFN- γ plus monoclonal anti-HuIFN- γ ; *b*, HuIFN- γ plus medium (control); *c*, ¹⁴C-labelled standard proteins.

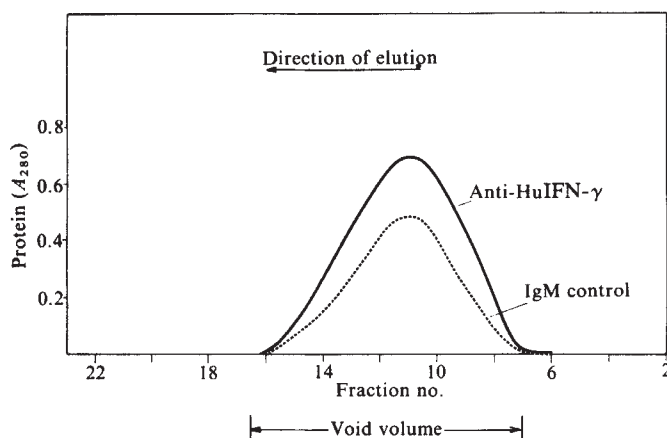


Fig. 1 Elution profile of monoclonal anti-HuIFN- γ on Sephadex G-200. Hybridoma supernatants (10 ml) containing $\sim 1\text{--}5 \mu\text{g ml}^{-1}$ of antibody were precipitated with 40% ammonium sulphate (ice-cold) and the precipitated proteins, after resuspension in $200 \mu\text{l}$ Tris-buffered saline pH 7.4, were further purified by filtration on a 10 ml Sephadex G-200 filtration column in Tris-buffered saline, pH 7.4. At this protein concentration no IgG protein peak was visible. The antibody eluted with the void volume and had an elution profile identical to that of an IgM control.

Iscoe's modified Dulbecco's medium (IMDM) supplemented with 20% fetal calf serum (FCS) and 1% methylcellulose. After 1 week, with 20% of hybridomas surviving, half of each individual culture was injected into the peritoneal cavity of separate BALB/c mice which had been pretreated with pristane for 1 week⁴. (The other half was cultured further in HAT medium.) One week after injection, ascites fluids were collected and the hybrids cultured in IMDM +10% FCS. Of 16 stable lines secreting monoclonal anti-HuIFN- γ , 15 were isolated from those cultures which had been passed through a mouse after fusion.

These hybridomas were subcloned in Greiner 'single clone' 24-well tissue culture plates in which each of the wells was divided into 16 subsquares. As the FO hybridomas grow attached to the well surface, these subdivided wells are a convenient tool for subcloning. Growing colonies were tested for specific anti-HuIFN- γ secretion in an antiviral neutralization test on human CCL 23 cells. Briefly, the surface of microtitre immunoassay plates were coated with rabbit anti-mouse immunoglobulin, and the supernatants of individual mouse hybridoma cultures were incubated in these plates for 16 h. After a final wash, 10 units of HuIFN- γ were added to each

individual well and the plates were incubated at 37 °C for 4 h. The remaining antiviral activity in each cup was then titrated in a viral RNA reduction assay using vesicular stomatitis virus on human CCL 23 cells. By thus testing the individual anti-IFN- γ hybridoma lines for constant antibody secretion over a period of 6 months, 16 stable mouse hybridoma lines secreting monoclonal anti-HuIFN- γ were established. So far, we have characterized the monoclonal antibody of one of these lines.

To identify the immunoglobulin species of this monoclonal anti-HuIFN- γ , the antibody was purified from hybridoma supernatants by ammonium sulphate (40%) precipitation followed by gel filtration on Sephadex G-200. The anti-HuIFN- γ eluted in the void volume with an elution profile identical to that of an IgM control (Fig. 1). The specificity of the purified monoclonal anti-HuIFN- γ was tested further in an enzyme-linked immunosorbent assay (ELISA) against HuIFN- γ which had been separated by SDS-polyacrylamide gel electrophoresis⁵ (omitting β -mercaptoethanol treatment of the HuIFN- γ sample) and electrophoretically transferred to a nitrocellulose sheet according to Towbin *et al.*⁶. Figure 2 shows

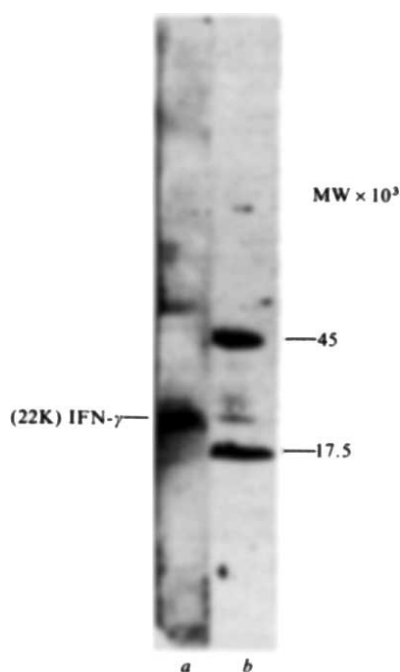


Fig. 3 ELISA using HuIFN- γ and monoclonal anti-HuIFN- γ . Partially purified HuIFN- γ which had been pretreated with SDS and β -mercaptoethanol was separated in parallel with standard proteins by SDS-polyacrylamide gel electrophoresis according to Laemmli⁵. After electrophoretic transfer of all proteins to a nitrocellulose sheet, the section containing the standard proteins was stained with amido black. An ELISA was performed on the section containing the HuIFN- γ as described in Fig. 2 legend. *a*, HuIFN- γ plus monoclonal anti-HuIFN- γ ; *b*, standard proteins stained with amido black.

that the monoclonal anti-HuIFN- γ specifically reacted with the HuIFN- γ preparation at a position corresponding to molecular weight (MW) 46,000 (46K). The same experiment was carried out after treatment of the IFN- γ preparation before SDS-polyacrylamide gel electrophoresis with SDS and β -mercaptoethanol to avoid possible dimerization of the HuIFN- γ molecule. In these conditions the HuIFN- γ blotted on to nitrocellulose reacted with monoclonal anti-HuIFN- γ at a position corresponding to MW 22,000 (Fig. 3). These results suggest that monoclonal anti-HuIFN- γ is specifically directed either against two different HuIFN- γ subspecies or against the dimer in one case (Fig. 2) and against the monomer in another (Fig. 3). However, the shifting of the immuno-reacting HuIFN- γ

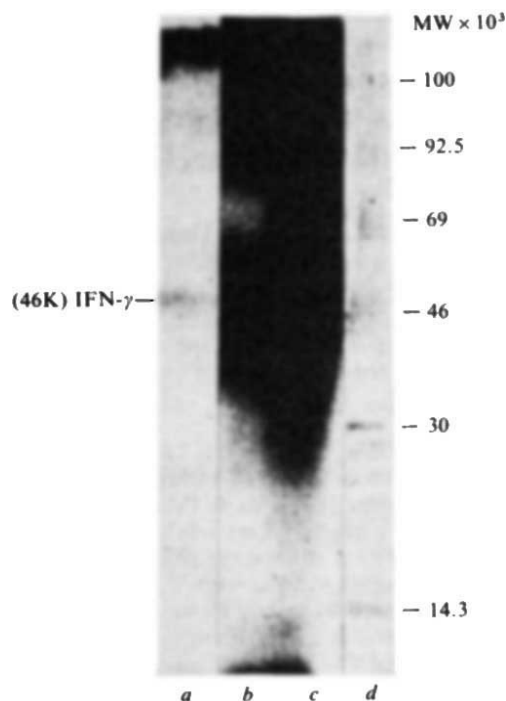


Fig. 4 Autoradiograph of a 15% SDS-polyacrylamide gel⁵ containing: *a*, immuno-purified ¹²⁵I-HuIFN- γ ; *b*, wash fraction; *c*, original ¹²⁵I-HuIFN- γ ; and *d*, ¹⁴C-labelled protein markers. Partially purified ¹²⁵I-IFN- γ which had been incubated first with monoclonal mouse anti-HuIFN- γ for 16 h and then with rabbit IgG antibody directed against mouse IgM (16 h) was precipitated with *S. aureus* protein A. The precipitate was washed four times with Tris-buffered saline, then the specifically bound ¹²⁵I-IFN- γ (pretreated with SDS, but not β -mercaptoethanol) was separated on an SDS-polyacrylamide gel⁵ in parallel with ¹⁴C-protein markers, the wash fraction and the original ¹²⁵I-HuIFN- γ sample.

band after β -mercaptoethanol treatment makes the latter possibility more likely.

In a further experiment, partially purified HuIFN- γ which had been labelled with ¹²⁵I according to Hunter *et al.*⁷ was incubated first with monoclonal mouse anti-HuIFN- γ and subsequently with a rabbit (IgG) anti-mouse IgM antibody. This complex was then precipitated with IgG-specific *Staphylococcus aureus* protein A. After several washes of the precipitate, the specifically bound ¹²⁵I-labelled IFN- γ was separated on an SDS-polyacrylamide gel according to Laemmli⁵ (without β -mercaptoethanol treatment) in parallel with ¹⁴C-labelled protein markers. Autoradiography of the dried gel (Fig. 4) showed that monoclonal anti-HuIFN- γ bound to *S. aureus* protein A, specifically precipitated ¹²⁵I-IFN- γ at 46,000 MW. Finally in an ELISA there was no cross-reaction of anti-HuIFN- γ with either HuIFN- β or HuIFN- α which had been electrophoretically transferred from an SDS-polyacrylamide gel on to a nitrocellulose sheet, nor did monoclonal anti-HuIFN- β cross-react with HuIFN- γ (data not shown).

HuIFN- β was a gift from Rentschler/Laupheim Co. and HuIFN- α was given by H. Mohr. M.L. is Bevoegdverklaard Navorsers of the Belgian National Fonds voor Wetenschappelyk Onderzoek.

Received 10 November 1981; accepted 21 January 1982.

1. Köhler, G. & Milstein, C. *Nature* **256**, 495–497 (1975); *Eur. J. Immun.* **6**, 511–519 (1976).
2. De Ley, M. *et al. Eur. J. Immun.* **10**, 877–883 (1980).
3. Hochkeppel, H. K., Menge, U. & Collins, J. *Eur. J. Biochem.* **118**, 437–442 (1981).
4. Staehelin, Th. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 1848–1852 (1981).
5. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
6. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
7. Hunter, W. M. & Greenwood, I. C. *Nature* **194**, 495–496 (1962).

Absence of significant H-2 and β_2 -microglobulin mRNA expression by mouse embryonal carcinoma cells

Dominique Morello*, Françoise Daniel,
Patricia Baldacci, Yvon Cayre,
Gabriel Gachelin & Philippe Kourilsky

Unité de Biologie Moléculaire du Gène, ER CNRS 201, SC
INSERM 20, Institut Pasteur, 75724 Paris Cedex 15, France

Class I histocompatibility antigens (H-2 antigens in the mouse) are 44,000-molecular weight cell-surface glycoproteins, the proper insertion of which depends on their noncovalent association with β_2 -microglobulin (β_2 m) (reviewed in ref. 1). They are involved in a variety of immune phenomena, particularly in self/non-self recognition (reviewed in ref. 2). H-2 antigens and β_2 m are expressed by nearly all adult somatic cells. The status of their expression on pre-implantation mouse embryos remains controversial^{3,4} although they are generally agreed to be serologically detectable at post-implantation stages⁵. H-2 and β_2 m genes are therefore developmentally regulated. H-2 antigens have not been found on murine embryonal carcinoma (EC) cells⁶, which are widely used as an alternative to studying normal early embryonic cells⁷⁻⁹, but they become detectable during *in vitro* differentiation of EC cells¹⁰⁻¹². During studies aimed at understanding the genetic mechanisms involved in the expression of H-2 antigens and β_2 m during differentiation, cDNAs reverse transcribed from mRNAs coding for H-2 or H-2-related heavy chains and β_2 m have been characterized¹³⁻¹⁶. H-2 and β_2 m cDNAs were used as probes to screen total poly(A)⁺ RNAs from various cell types. We report here that trace amounts of poly(A)⁺ RNA from EC cells were found to hybridize with these probes, finding which contrasts with results obtained on poly(A)⁺ RNA prepared from differentiated cells. These results, taken together with the absence of major rearrangements in H-2 and β_2 m genes during differentiation, suggest that H-2 and β_2 m expression is likely to be controlled at the level of transcription.

Figure 1 legend describes the probes used in the study. Poly(A)⁺ RNAs were extracted from liver, spleen and kidney of DBA/2 (H-2^d) and 129/Sv (H-2^b) mice. (These animals are respectively syngeneic to SL2 tumour cells, from which the mRNAs used for cloning the probes have been isolated¹³, and to most of the EC lines used in the present study¹⁷). The poly(A)⁺ RNAs were electrophoresed in agarose gels, transferred to nitrocellulose sheets and finally hybridized with ³²P-labelled probes¹⁸. Probes H-2A, B and C revealed the same unique band, corresponding to RNAs of ~1,800 nucleotides (Fig. 2, lanes b, c, results shown for H-2B probe), a size in good agreement with that estimated (17S) from sedimentation through sucrose gradient¹⁹. In the various H-2 cDNA clones studied so far^{20,21}, the 3'-untranslated region is ~480 nucleotides long, excluding the poly(A) tail. The coding region, estimated from the known amino acid sequence of H-2K^b (ref. 22), must be about 1,000 nucleotides long; assuming the poly(A) tail does not exceed 100 nucleotides, the length of the 5'-untranslated region in H-2 mRNAs should be 100–300 nucleotides. We have detected no size difference nor any heterogeneity in the two mouse strains or in the various tissues tested (Fig. 2, lanes b, c and data not shown). The absence of heterogeneity implies that members of the H-2 multigene family, when expressed, yield mRNAs of an approximately similar size (in this size range, we could only detect a difference >100 nucleotides). Finally, if non-H-2 mRNAs are expressed and carry the reiterated sequence detected by probe C, they must fall within the same size ranges as H-2 mRNAs. Hybridization with the

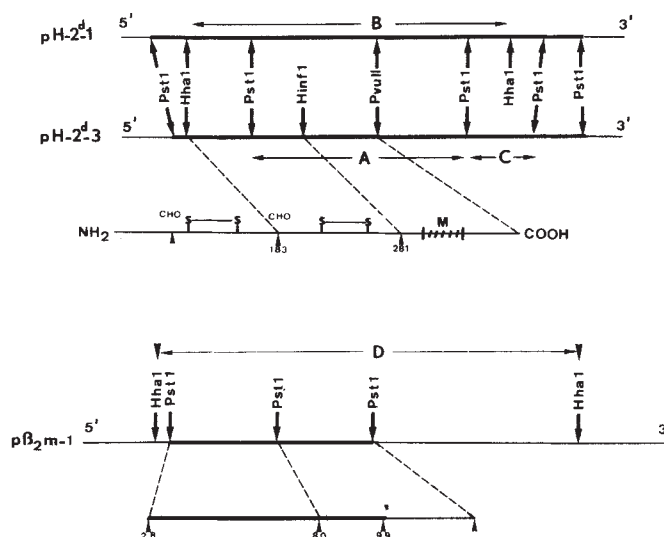


Fig. 1 Top: The restriction maps of the cDNA inserts of the recombinant plasmids pH-2^d-3 and pH-2^d-1 (refs 20, 24) are presented in correspondence with the linear structure of an H-2 antigen molecule. M indicates the membrane spanning region with the external part of the molecule on the NH₂-terminal side and the cytoplasmic fragment on the COOH-terminal side. A, B and C represent restriction fragments used as radiolabelled probes in experiments described in the text. Probe H-2A is the 578-base pair (bp) *Pst*I-*Pst*I of pH-2^d-3²⁰. From its 5' end, it extends over the COOH-terminal part of an H-2 molecule (nucleotides 180–498, *Pvu*II site), corresponding to part of the third external domain, transmembrane and cytoplasmic regions; the probe extends downstream over the first 260 bp of the noncoding region. This fragment does not contain a region hybridizing with a sequence highly repeated in the genome²⁴. Probe H-2B is the 938-bp *Hha*-*Hha* fragment of pH-2^d-1¹³. The 5' part (up to the *Pvu*II site) encodes for a nearly complete third extracellular domain, transmembrane and cytoplasmic regions²⁰. The 327-bp noncoding region downstream contains at its 3' end a stretch of 20 nucleotides belonging to the *Alu* type II sequence³⁰. Probe H-2C is the 160-bp *Pst*I-*Pst*I fragment of pH-2^d-3. Its 3' end includes the *Alu* type II sequence. This sequence is repeated a minimum of 40,000 times in the haploid genome of the mouse^{14,24}. Bottom: The restriction map of the cDNA insert of the recombinant plasmid p β_2 m-1 is presented in correspondence with the linear structure of the β_2 m molecule. This clone has been obtained by screening a mouse cDNA library with a recently cloned human β_2 m cDNA probe¹⁶. p β_2 m-1 includes most of the sequence coding for β_2 m and about 100 additional basepairs of the 3'-untranslated region. The restriction fragment D was used as the β_2 m probe.

β_2 m probe reveals two bands of similar intensities, migrating approximately as 1,000 and 800 nucleotides in both DBA/2 (Fig. 2, lane d, and Fig. 3A, lanes e, f) and 129/Sv (Fig. 3A, lane g) poly(A)⁺ RNAs. These two bands could correspond to either the same coding sequence associated with different untranslated regions or to products of two different genes.

Poly(A)⁺ RNAs were extracted from several EC lines, F9, PCC3 and PCC4-aza¹² and analysed as above. Using probes prepared at the usual specific radioactivity ($2-3 \times 10^8$ c.p.m. per μ g), no poly(A)⁺ RNA was found to hybridize with the H-2B, C and the β_2 m probes (Fig. 3A, lanes b–d; results shown with probes B and D even with autoradiographic exposure for 2 weeks (data not shown)). These poly(A)⁺ RNAs were considered as undegraded because they hybridized strongly with an actin probe, yielding an intense band with an apparent size of 2,000–2,100 nucleotides and no smear (Fig. 3B). When H-2 and β_2 m probes, displaying a much higher specific radioactivity (2×10^9 c.p.m. per μ g) were used, faint bands were observed on poly(A)⁺ RNAs extracted from EC cells, at the same location as in adult cells. No additional bands were observed on the latter control (Fig. 3C, lanes a, c). By comparison of the intensities of the bands obtained with similar amounts of poly(A)⁺ RNA prepared from EC and adult cells, we conclude that the former contained about 100 times less H-2- and β_2 m-specific mRNA than the latter. This result can be accounted for by either a normal expression in about 1% of the cells or by a 1% expression in most of them. Serological data¹² favour the first hypothesis.

Hybridization to H-2 and β_2m probes was carried out using poly(A)⁺ extracted from differentiated cells derived from EC lines by *in vitro* differentiation. We observed the same pattern as with poly(A)⁺ extracted from adult organs: one band migrating as ~1,800 nucleotides hybridizing with H-2 probe and two bands migrating approximately as 1,000 and 800 nucleotides hybridizing with β_2m probe. The intensity of the bands was dependent on the relative amount of H-2 antigens expressed by the cells tested: faint bands are observed with mRNAs prepared from 3/A/1-D-3 (ref. 17), a fibroblastic line derived from PCC3 exhibiting a low amount of H-2 antigens (data not shown). The intensity of the bands observed with mRNAs prepared from C17-S1-D-T 984 (ref. 23), a muscle cell line isolated from a contractile zone of cultured cells derived from a C3H teratocarcinoma, indicates that there is about the same amount of specific H-2 and β_2m mRNA in these cells as in adult cells (Fig. 3C, lanes b, c), confirming the immunofluorescence data.

These results suggest that there is about 1% of H-2 and β_2m poly(A)⁺ RNA in EC cells as compared with adult cells. Interestingly, this applies to all members of the H-2 multigene family^{14,24}, as well as to any other mRNA which could contain the reiterated sequence detected with probe H-2C. Moreover, our results rule out the possibility that H-2 heavy chains are synthesized in EC cells and remain undetected at the cell surface through lack of β_2m , as reported in several cell lines such as the human Daudi cells²⁵.

To determine whether major rearrangements at the DNA level would account for the absence of H-2 and β_2m expression in EC cells, we have carried out Southern blots using DNAs extracted from F9, PCC4 EC cells and from 129/Sv liver. DNAs were digested with *Bam*HI and *Eco*RI, run in 0.7% agarose gels, transferred to nitrocellulose sheets and hybridized with H-2B and β_2m probes. The multi-band patterns obtained with probe H-2B on EC and liver DNAs are identical as far as the sizes of the bands are concerned; however, differences in the intensity of several bands do exist (Fig. 4). Two major bands were detected when using the β_2m probe, and their sizes did not change on differentiation (data not shown). These findings confirm the absence of major DNA rearrangements in and around β_2m and H-2 genes during differentiation²⁶.

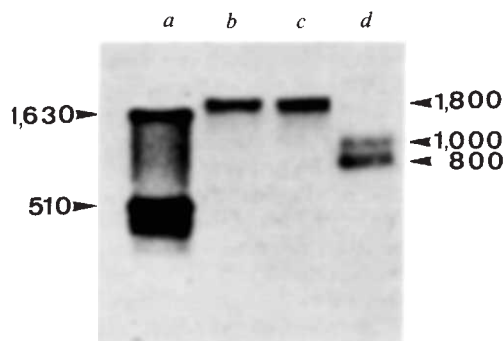


Fig. 2 Poly(A)⁺ RNAs were extracted from adult organs, precipitated with ethanol and isolated by binding twice to oligo(dT) cellulose as described elsewhere³¹. 5 μ g of each preparation were subjected to electrophoresis in a 1.1% agarose gel and transferred to nitrocellulose sheets, also as described elsewhere¹⁸. The sheets were prehybridized, hybridized to ³²P-radiolabelled restriction fragments. A, B, C and D restriction fragments (Fig. 1) were polymerized using T4 DNA ligase (Boehringer Mannheim) and then ³²P radiolabelled by nick-translation³² to a specific activity of 2–3 $\times 10^8$ c.p.m. per μ g. Each nitrocellulose sheet was hybridized with 10⁷ c.p.m. of each probe for 24 h and then washed as described previously¹⁸. The sheets were then autoradiographed on flash-activated Kodak X-Omat films for 2–14 days depending on experiments, at –70 °C with intensifying screens. Lane a, 2 ng of pBR322 digested by *Hinf*I were hybridized with a pBR322 probe. Lanes b, c, 5 μ g of poly(A)⁺ RNA respectively extracted from DBA/2 liver and from 129/Sv liver, spleen and kidney, were hybridized with probe B. Lane d, 5 μ g of poly(A)⁺ RNA extracted from liver DBA/2 were hybridized with probe D. (Autoradiographic exposure time 48 h.) In our experimental conditions, 18S rRNA migrates slightly above 1,800 bp (a location corresponding to a size of 1,920 bp DNA).

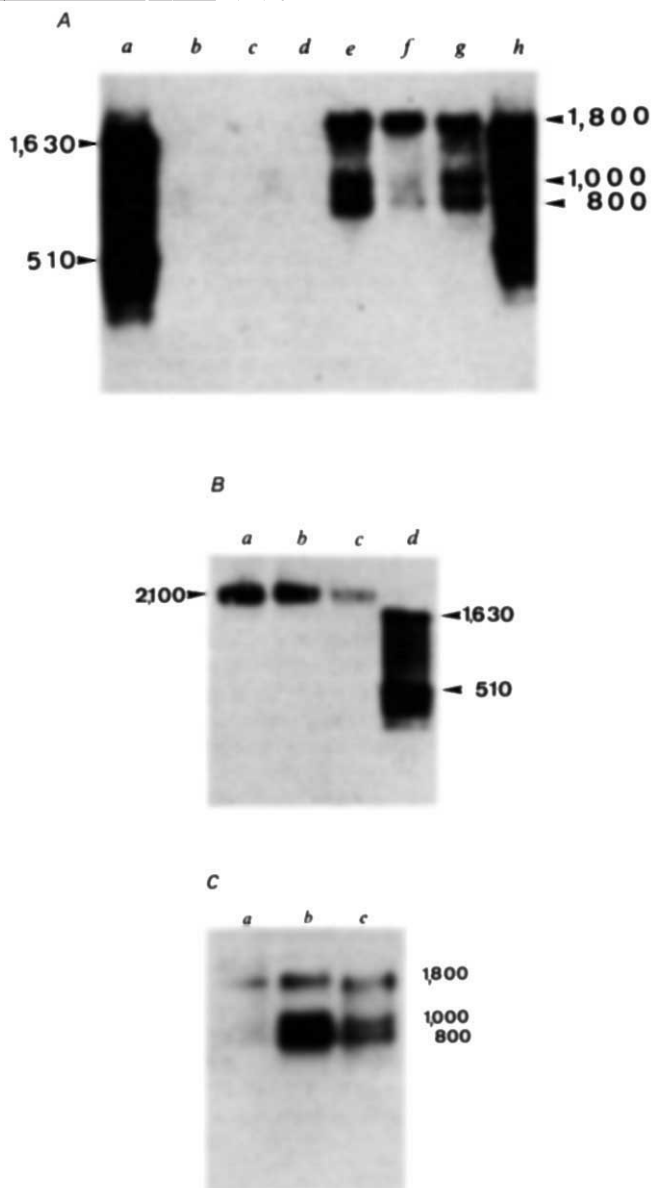


Fig. 3 A, Poly(A)⁺ RNAs were extracted from EC cell lines and used for Northern blotting as described in Fig. 2 legend. Lanes a, h, 10 ng of pBR322 digested by *Hinf*I were hybridized with a pBR322 probe. Lanes b–d, 5 μ g of poly(A)⁺ RNAs respectively extracted from PCC4-aza, F9 and PCC3 EC cell lines, were hybridized with probes B and D. Lanes e–g, 5 μ g of poly(A)⁺ respectively extracted from DBA/2 liver, DBA/2 spleen and 129/Sv liver, spleen and kidney, were hybridized with probes B and D. (Autoradiographic exposure time 48 h.) B, Lanes a–c, 5 μ g of poly(A)⁺ RNAs, respectively extracted from PCC3, F9 and PCC4-aza, were hybridized with an actin probe. Lane d, 2 ng of pBR322 digested by *Hinf*I were hybridized with a pBR322 probe. (Autoradiographic exposure time 48 h.) C, Lane a, 10 μ g of poly(A)⁺ RNA extracted from PCC4-aza EC cells. Lane b, 4 μ g of poly(A)⁺ extracted from C17-S1-D-T 984 (muscle cell line). Lane c, 1 μ g of poly(A)⁺ RNA extracted from 129/Sv liver, spleen and kidney. The nitrocellulose filter was hybridized simultaneously with B and D probes. (Autoradiographic exposure 24 h.)

It thus seems likely that H-2, H-2-related genes and the β_2m gene(s), taken here as developmentally regulated genes, are primarily controlled at the level of transcription. These genes belong to different chromosomes (H-2, TL, Qa genes, lie on chromosome 17, β_2m probably maps on chromosome 2²⁷; the locations of H-2-related genes are not precisely known). Therefore, no common *cis*-regulatory process could be postulated on a topological basis.

Croce *et al.*²⁸ published similar but not identical results on F9 EC cells transformed with a plasmid carrying herpes simplex virus type 1 TK gene and SV40 genes²⁹. The expression of H-2

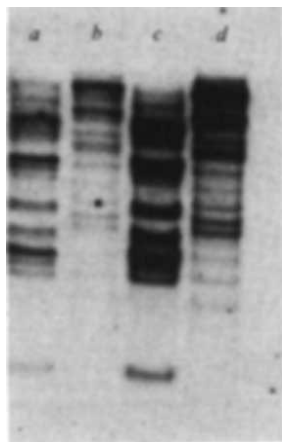


Fig. 4 The high molecular weight genomic DNA was extracted and purified as described previously²⁴ from 129/Sv mouse liver and from F9 EC cells. DNA samples (20 µg) were digested by the appropriate restriction enzymes (*Bam*HI or *Eco*RI), subjected to electrophoresis in a 0.7% agarose gel, transferred to nitrocellulose sheets and hybridized as described previously²⁴ with probe B. Lanes a–b, 129/Sv DNA digested with *Bam*HI and *Eco*RI; lanes c–d, F9 DNA digested with *Bam*HI and *Eco*RI.

and SV40 antigens was induced by treatment of the cells by retinoic acid. Their results are identical to ours concerning the absence of specific H-2 and β_2m mRNAs in nondifferentiated cells, but their results on differentiated cells differ from ours. Croce *et al.* obtained a hybridization to a smear and not to distinct bands. This smear extends to much lower molecular weight mRNAs, suggesting either degradation or alteration of a normal message due either to the expression of the inserted plasmid or treatment with retinoic acid. In contrast, clonal lines of differentiated derivatives of EC cells display, in our studies, the same pattern as normal differentiated cells. Thus, the pattern observed by Croce *et al.* might reveal changes in a normal gene expression caused by a non-physiological differentiator agent or interference with viral genes.

The actine probe was made available to us by F. Brégégère and A. Minty. We thank H. Jacob for many helpful discussions and the gift of EC cells, D. Rocancourt for technical assistance and M. Caravatti for his gift of C17-S1-D-T 984 poly(A)⁺ RNA. This work was supported by grants from CNRS (ER 201, ATP 4246 and 5039), INSERM (SC 20, CRAT 72.79.104) and DGRST (ATP 79.7.0500 and 78.7.2931). The work was carried out in accordance with the French Guidelines for Recombinant Research.

Received 26 October 1981; accepted 26 January 1982.

- Ploegh, H. L., Orr, H. T. & Strominger, J. L. *Cell* **24**, 287–299 (1981).
- Klein, J., Juretic, A., Baxevas, C. N. & Nagy, Z. A. *Nature* **291**, 455–460 (1981).
- Webb, C. G., Gall, E. & Edelman, G. M. *J. exp. Med.* **146**, 923–932 (1977).
- Sawicki, J. A., Magnuson, T. & Epstein, C. J. *Nature* **294**, 450–451 (1981).
- Buc-Caron, M. H., Condamine, H. & Jacob, F. *J. Embryol. exp. Morph.* **47**, 149–160 (1978).
- Morello, D. *Transplantation* **26**, 119–125 (1978).
- Strickland, S. *Cell* **24**, 277–278 (1981).
- Evans, M. J. & Kaufman, M. H. *Nature* **292**, 154–156 (1981).
- Martin, G. R. *Science* **209**, 768–776 (1980).
- Nicolas, J. F., Dubois, P., Jakob, H., Gaillard, J. & Jacob, F. *Annls Microbiol. Inst. Pasteur, Paris* **A126**, 3–22 (1975).
- Stern, P. L., Martin, G. R. & Evans, M. J. *Cell* **6**, 455–465 (1975).
- Gachelin, G. *Biochim. biophys. Acta* **516**, 27–60 (1978).
- Kvist, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2772–2776 (1981).
- Steinmetz, M. *et al. Cell* **24**, 125–134 (1981).
- Parnes, J. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2213–2217 (1981).
- Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6613–6617 (1981).
- Nicolas, J. F. *et al. Cancer Res.* **36**, 4229–4231 (1976).
- Thomas, S. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
- Dobberstein, B., Garoff, H., Warren, G. & Robinson, P. *Cell* **17**, 759–769 (1979).
- Brégégère, F. *et al. Nature* **292**, 78–81 (1981).
- Lalanne, J. L. *et al. Nucleic Acids Res.* (in the press).
- Colligan, J. F., Kindt, T. J., Uehara, H., Martinko, J. & Nathenson, S. G. *Nature* **291**, 35–39 (1981).
- Jakob, H. *et al. Expl Cell Res.* **114**, 403–408 (1978).
- Cami, B., Brégégère, F., Abastado, J. P. & Kourilsky, P. *Nature* **291**, 673–675 (1981).
- Sege, K., Rask, L. & Peterson, P. R. *Biochemistry* **20**, 4523–4529 (1981).
- Steinmetz, M. *et al. Cell* **25**, 683–692 (1981).
- Michaelson, J. *Immunogenetics* **13**, 167–171 (1981).
- Croce, C. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 5754–5758 (1981).
- Linnenbach, A., Huebner, K. & Croce, C. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4875–4879 (1980).
- Haynes, S. R. & Jelinek, W. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6130–6134 (1981).
- Aufray, C. & Rougeon, F. *Eur. J. Biochem.* **107**, 303–314 (1980).
- Niatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).

Nuclear localization and DNA binding of the transforming gene product of avian myelocytomatosis virus

Peter Donner, Irene Greiser-Wilke & Karin Moelling

Max-Planck-Institut für Molekulare Genetik, Berlin-Dahlem, 1 Berlin 33, Ihnestr. 63-73, FRG

Oncornaviruses transform cells either directly through a virus-coded oncogene¹ or, if they lack such a gene, indirectly by promoter insertion^{2–4} into the cellular genome and activation of a cellular gene. Both transformation mechanisms ultimately result in abnormally high expression of normal genes. Recently, the activated cellular gene in certain lymphomas⁴ has been shown to be homologous to the oncogene of an acute avian leukaemia virus, MC29. In MC29 the oncogene is fused to the viral structural gene, *gag*. The product of this fused gene is a protein of molecular weight 110,000 (110 K)^{5–7}, designated p110^{gag-myc}. We have characterized this protein by using monoclonal antibodies against p19, the N-terminal portion of the *gag-myc* fusion or 110 K protein and purified it 3,700-fold by immune affinity column chromatography. Immunofluorescence studies and cell fractionation of MC29-transformed fibroblasts indicate that the 110 K protein is predominantly located in the nucleus. Moreover, the purified protein binds to double-stranded DNA. These properties may be related to the role of the protein in transformation.

The MC29-transformed quail embryo fibroblast cell line, MC29-Q8-NP, does not produce virus particles but expresses the transforming protein *gag-myc*, which is fused to the viral structural *gag* proteins p19 and p27⁵. The monoclonal antibody against p19⁸ allowed immunoprecipitation of the 110 K protein (Fig. 1A). Superinfection of this cell line with a helper virus, Rous-associated virus RAV-60, resulted in synthesis of viral polyprotein precursors which were about five times more abundant than the 110K protein. RAV-60-infected quail fibroblasts exhibited viral polyprotein precursors but no 110K protein (Fig. 1A). Indirect immunofluorescence using the IgG of monoclonal anti-p19 and fluorescein-conjugated anti-mouse IgG⁹ revealed that the MC29-Q8-NP cells showed a strong nuclear fluorescence (Fig. 2a, d) while the helper virus-superinfected cells showed a predominantly cytoplasmic staining (Fig. 2b). Normal quail embryo fibroblasts did not give rise to a positive fluorescence (Fig. 2c), while predominantly cytoplasmic fluorescence was seen in quail embryo fibroblasts infected with RAV-60 (Fig. 2e) and in the Rous sarcoma virus-transformed cell line R(-)Q, which expresses the transforming protein pp60^{src} and viral polyprotein precursors¹⁰ (Fig. 2f). Therefore, the nuclear fluorescence of MC29-transformed quail fibroblasts seems to be specific for transformation by MC29. All other known oncornaviral transforming proteins were shown to be localized mainly in the cytoplasmic membrane¹¹.

To confirm our observation, a cell fractionation from radioactively labelled MC29-Q8-NP cells and RAV-60-superinfected MC29-Q8-NP cells was performed. Nuclei were isolated by hypotonic shock and Dounce homogenization and subsequently stripped by treatment with nonionic detergent¹², a procedure which leaves behind a crude chromatin preparation. The soluble fractions were combined and contained endoplasmic reticulum, organelles, cytoplasmic and nuclear membranes and polysomes. The stripped nuclei contained 70% of the 110K protein, while the residual 30% were present in the soluble fraction, which also contained >90% of each of the viral proteins Pr180^{gag-pol}, Pr76^{gag}, p27, Pr92^{env} and gp85. These values were derived from the radioactivity eluted from the various protein bands of the gel shown in Fig. 1B and C and corrected

for differences in volumes. The 110K protein was only detectable in the soluble fraction if 10 times more radioactive material was used (Fig. 1B).

For further analysis the 110K protein was purified from ^{35}S -methionine-labelled MC29-Q8-NP cells by an immune affinity column containing protein A-Sepharose to which monoclonal anti-p19 IgG was covalently coupled in conditions which link the constant region of the IgG molecule to the solid phase¹³. The 110K protein was eluted from the column with a buffer of low pH (pH 2) and salt (Fig. 3a), conditions which have been applied to purify biologically active human leukocyte interferon¹⁴. Comparison of the radioactive material eluted from the immune affinity column with the total radioactive input indicated a 3,700-fold purification. The eluted 110K protein was precipitable by anti-p19 serum, indicating that it was still antigenically active (Fig. 1, slots a, b). The p19-containing viral polyprotein precursors in virus-producing cells were isolated by the identical purification procedure from Rous sarcoma virus-infected chicken embryo fibroblasts (Fig. 3b). Many p19-related polyproteins were eluted, the majority of which was represented by Pr76^{gag}, which served as a control for nucleic acid-binding studies described below.

Because the localization of the 110K protein in the nucleus raised the question of whether it interacted with cellular DNA, the purified protein was analysed for its nucleic acid-binding properties *in vitro*. DNA of a defined size class was preparatively isolated from avian fibroblasts¹⁵, mixed with the ^{35}S -methionine-labelled purified 110K protein and sedimented in a glycerol gradient. In these conditions the 110K protein comigrated with the DNA, whereas without DNA it remained on top of a gradient run as a control (Fig. 4a). To determine whether the nucleic acid-binding property of the 110K protein was attributable to the gag or the transforming protein portion, Pr76^{gag} and p19-related polyproteins were analysed in the same way. These proteins did not bind to the DNA (Fig. 4b), indicating that it is the transformation-specific moiety of the 110K protein which interacts with the DNA.

The nucleic acid-binding property of the 110K protein was also demonstrated by using radioactively labelled DNA to measure the ability of the 110K protein to retain DNA on a nitrocellulose filter in a filter binding assay¹⁶. The 110K but not the Pr76^{gag} protein allowed the binding of DNA quite efficiently (Fig. 4c). Use of ^{35}S -methionine-labelled proteins enabled us to control the linearity of binding of the protein to

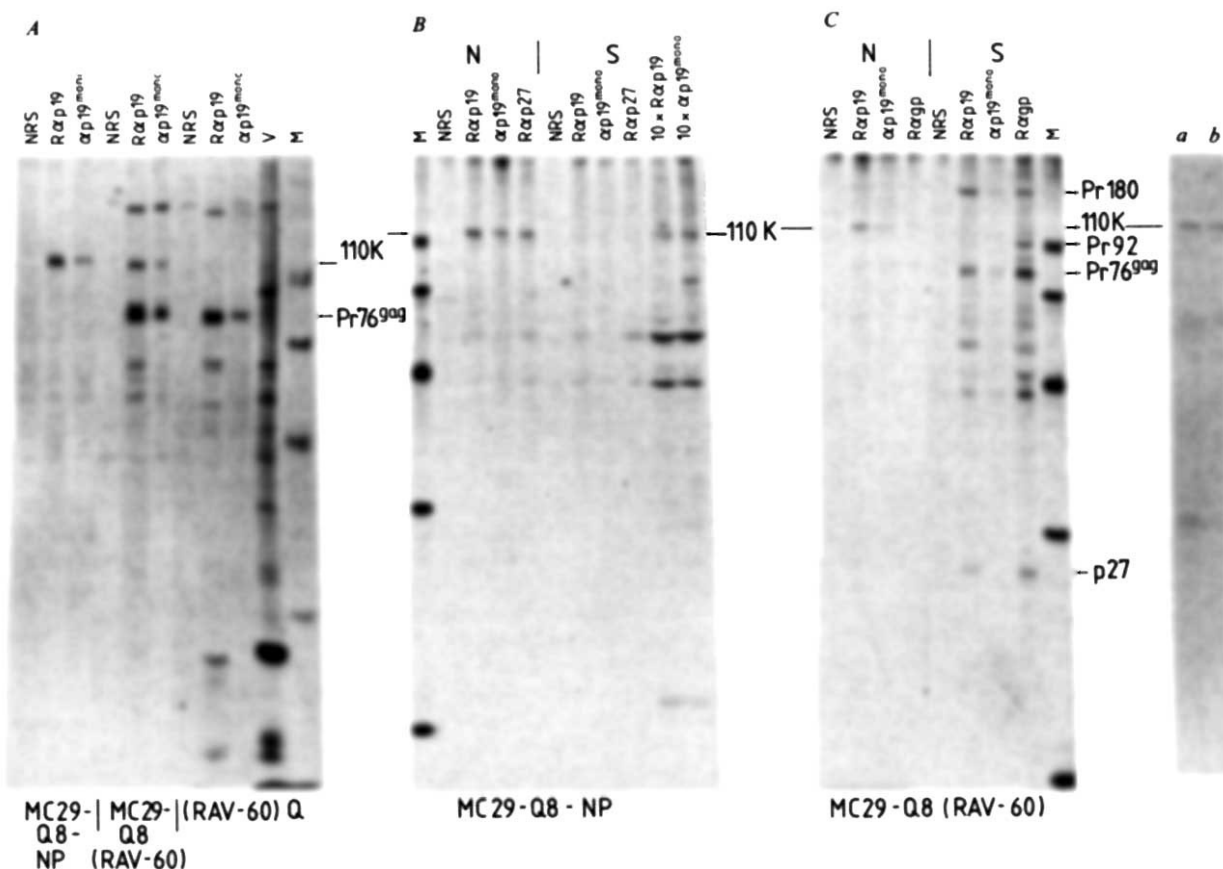


Fig. 1 Distribution of 110K protein in cells and subcellular fractions. **A**, MC29-Q8-NP cells, MC29-Q8 cells superinfected with RAV-60 (MC29-Q8 (RAV-60)) and RAV-60-infected quail fibroblasts ((RAV-60)Q) (3×10^6 cells each) were labelled with ^{35}S -methionine ($250 \mu\text{Ci ml}^{-1}$) for 4 h, lysed and treated with anti sera and *Staphylococcus aureus*²¹ for indirect immunoprecipitation. The precipitates were analysed on polyacrylamide slab gels and processed for autoradiography²⁰. The sera used were: NRS; normal rabbit serum; Rαp19, rabbit anti-p19; αp19^{mono}, purified IgG of monoclonal antibodies against p19; 5 μl of each serum was used and 10 μg of the IgG. V represents ^{35}S -methionine-labelled RAV-60 virus and M a set of ^{14}C -labelled marker proteins (Amersham) (from top) molecular weights of 92.5K (upper band of doublet), 69K, 46K, 30K and 14K. **B**, Isolation of stripped nuclei was performed according to published procedures¹². MC29-Q8-NP cells (1.5×10^7) were radioactively labelled with $250 \mu\text{Ci ml}^{-1}$ of ^{35}S -methionine²⁰ for 4 h and processed immediately. The plates were washed three times with ice-cold phosphate-buffered saline (PBS, 17 mM Na_2HPO_4 , 2.6 mM KH_2PO_4 pH 7.4, 120 mM NaCl) and treated with hypotonic buffer (10 mM Tris-HCl pH 7.3, 10 mM NaCl, 1.5 mM MgCl_2) for 40 min at 4 °C for lysis. Subsequently the cells were homogenized in a tightly fitting Dounce homogenizer with 10–20 strokes. Release of nuclei was followed by phase contrast microscopy. The nuclei were pelleted at 3,000 r.p.m. for 5 min, and the pellet was resuspended in hypotonic buffer and treated with 1% NP40 and 0.5% desoxycholate (DOC) to strip the nuclei. After mixing on a Vortex for 4 s the stripped nuclei (N) were pelleted (3,000 r.p.m., 5 min). The soluble supernatants (S) were combined and brought up to 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 0.1% SDS, 1% DOC, 1% Triton X-100 (RIPA buffer). The stripped nuclei were suspended in 1 ml RIPA buffer. After incubation for 15 min at 0 °C, samples were centrifuged (10,000g for 30 min). The equivalent of 1×10^6 c.p.m. of the supernatants were treated for immunoprecipitation as described in **A**; for nomenclature also see **A**. Rαp27, rabbit serum against p27; 10 $\times$, immunoprecipitation of 10 times more soluble fraction with anti-p19 sera. **C**, Isolation of subcellular fractions from MC29-Q8-NP cells superinfected with RAV-60 was performed as described in **B**. For nomenclature see **A** and **B**. Rαgp, rabbit antiserum against gp85 (which precipitates Pr92^{env} in addition to gp85 and also recognizes p27 and related polyproteins). Slot a shows the immune precipitate of a 110K-containing fraction (400 μl) from the immune affinity column (Fig. 3a) by anti-p19 monoclonal IgG and *S. aureus*. Slot b shows 50 μl of the same fraction applied directly to the gel (12.5%).

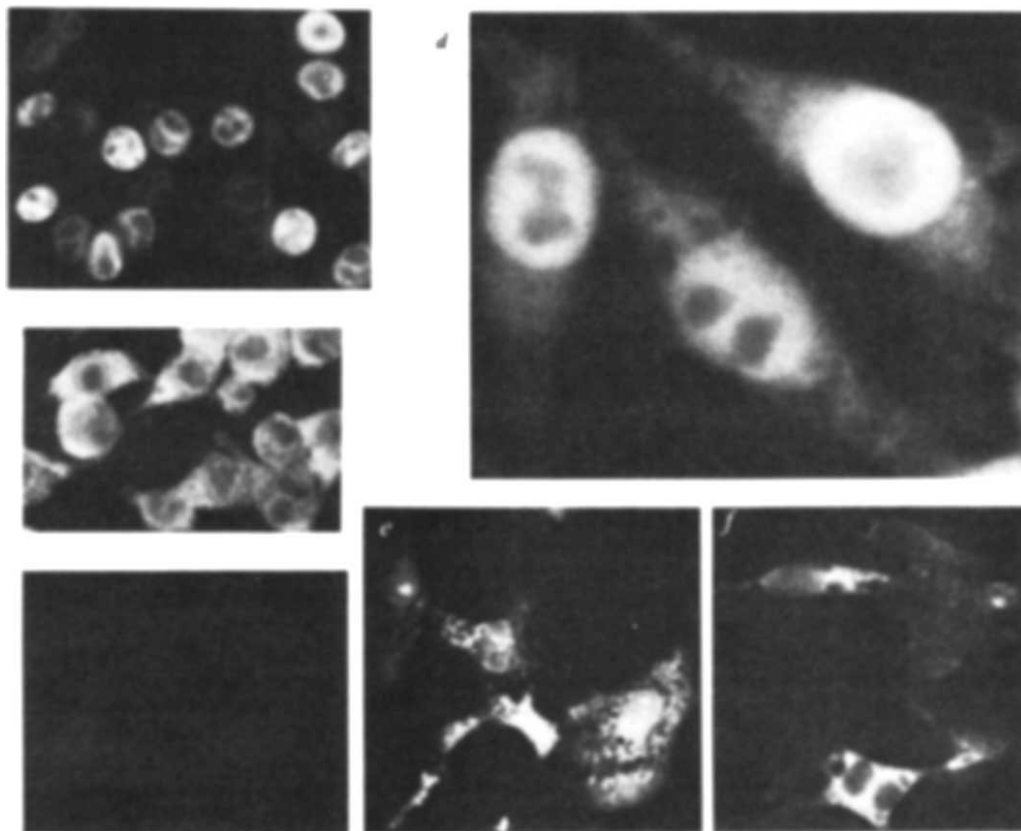


Fig. 2 Immunofluorescence by i-p19 antibody of transformed and superinfected cells. Indirect immunofluorescence was performed using monoclonal antibodies against p19 and fluorescein-labelled second antibody on the following cells: *a*, *d*, MC29-Q8-NP; *b*, MC29-Q8-NP superinfected with RAV-60; *c*, normal quail embryo fibroblasts; *e*, RAV-60-infected quail embryo fibroblasts; *f*, a quail cell line transformed by a defective Rous sarcoma virus, R(-)Q10. For fluorescence analysis cells were grown on a coverslip to a subconfluent stage. They were washed with PBS, fixed in 2.7% paraformaldehyde solution for 30 min at room temperature and rinsed with PBS. The cells were permeabilized in 50% ice-cold acetone/PBS solution for 2 min, 100% acetone for 5 min and 40% acetone/PBS solution for 2 min, rinsed in PBS and treated with monoclonal anti-p19 IgG solution ($1 \mu\text{g ml}^{-1}$) for 30 min at 37 °C (for the isolation of IgG see Fig. 3 legend). Excess antibody is removed by rinsing with PBS. Subsequently, fluorescein isothiocyanate-conjugated anti-mouse IgG was added. Incubation is for 30 min at 37 °C and was followed by three rinses with PBS. $\times 1,200$, except for *d*: $\times 3,000$.

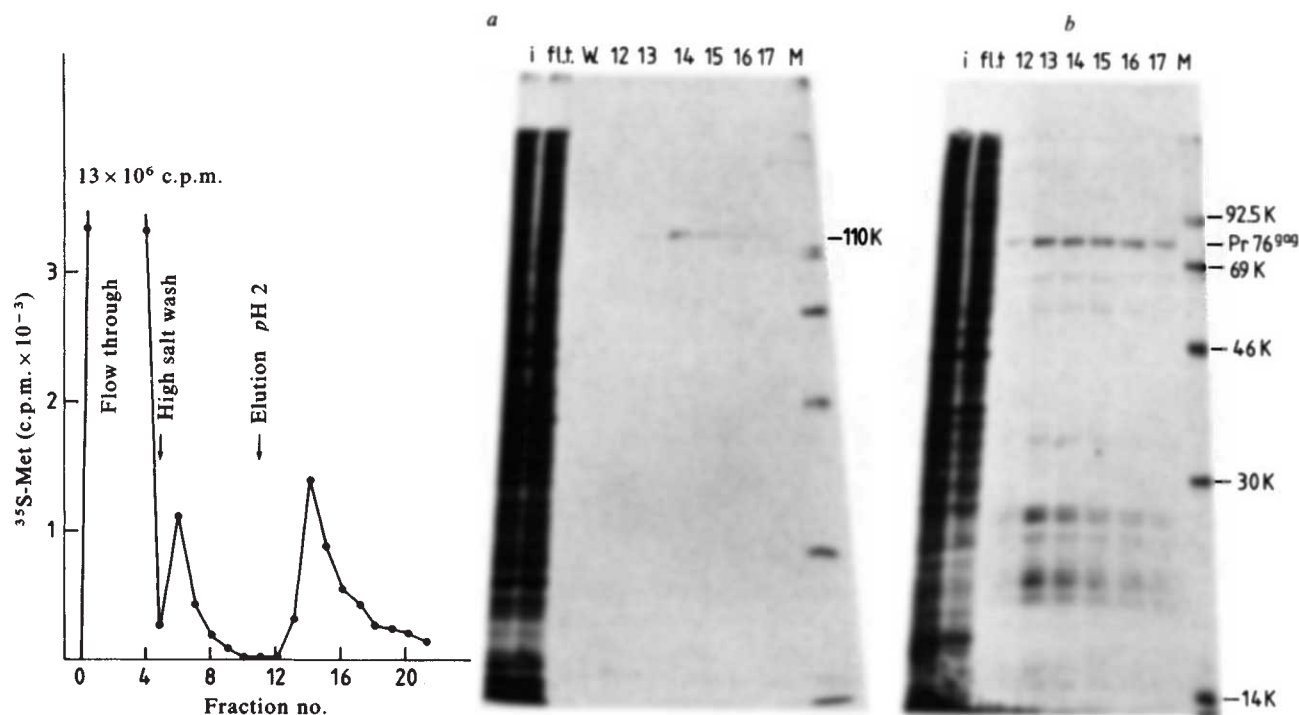
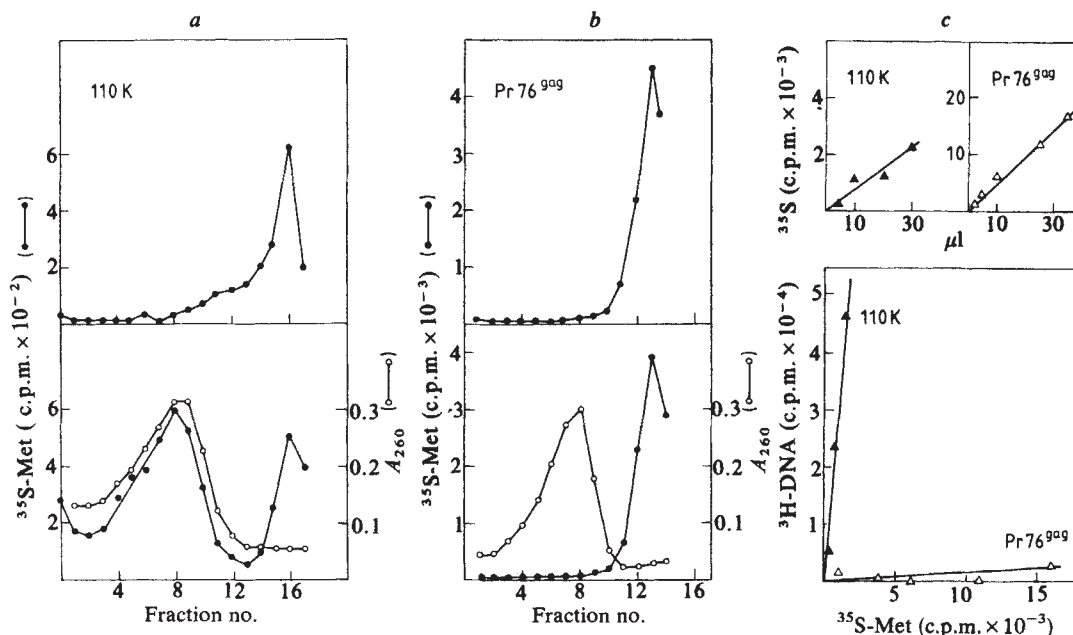


Fig. 3 Isolation of 110K protein from MC29-Q8-NP cells and Pr76^{gag} from virus-producing chicken fibroblasts by immune affinity chromatography. *a*, IgG from monoclonal anti-p19 was purified by protein A-Sepharose column chromatography (Pharmacia) from hybridoma supernatant and eluted with 4 M MgCl₂. The immunoabsorbent column was prepared according to Gersten and Marchalonis¹³: 7 mg of purified anti-p19 IgG were covalently coupled to 1 ml of protein A-Sepharose CL-4B by dimethylsuberimidate (30 mg) in 0.1 M borate buffer, pH 8.0. Then the column was pretreated successively with 1 ml each of PBS, PBS plus bovine serum albumin (1 mg ml^{-1}), PBS, a lysate of 10^8 normal quail embryo fibroblasts, and PBS. For preparation of the cell lysate, 5×10^7 MC29-Q8-NP cells were labelled with $250 \mu\text{Ci ml}^{-1}$ of ^{35}S -methionine for 4 h (ref. 20), lysed with 10 ml RIPA buffer in the presence of trasylol (Bayer) and centrifuged at 10,000 r.p.m. at 4 °C for 30 min. The supernatant (12 ml containing 6.4×10^8 c.p.m.) was then applied to the immune affinity column (200 μl bed volume). The column was washed successively with 5 ml RIPA buffer, 5 ml high-salt buffer (20 mM Tris-HCl pH 7.5, 1 M NaCl, 0.5% NP40) and 2 ml PBS before the protein was eluted with a buffer containing 0.1 M citric acid pH 2, 300 mM NaCl, 50% ethylene glycol, 0.1% Triton X-100. Fractions of 400 μl were collected, 20 μl of each fraction was counted in a scintillation counter and 50 μl applied to a SDS-polyacrylamide slab gel (10%). In total, 1.8×10^5 c.p.m. were eluted, indicating a 3,700-fold enrichment. The 110K protein-containing fractions were diluted and dialysed overnight against 50 mM phosphate buffer pH 6.2, 100 mM NaCl, 1 mM dithiothreitol and 10% glycerol. *i*, 5 μl of the input (12 ml); *ft.*, 5 μl of the flow-through; *W*, 5 μl of the pooled wash (10 ml); 12–17, eluted fractions; *M*, ^{14}C -labelled marker proteins (see Fig. 1 legend). *b*, The identical procedure was performed with Rous sarcoma virus-infected chicken embryo fibroblasts. The p19-related proteins eluted from the column are shown after gel electrophoresis and autoradiography. Pr76^{gag} represents the majority of the eluted p19-related proteins.

Fig. 4 Interaction of the 110K protein with DNA. *a*, DNA was extracted from chicken embryo fibroblasts¹⁵ and sheared by pressing it six times through a 21-gauge needle. The DNA was centrifuged through a 10–30% glycerol gradient in a SW41 rotor at 40,000 r.p.m. for 5 h, and the peak fraction was precipitated with ethanol. ³⁵S-methionine-labelled dialysed 110K protein from the immune affinity column (fractions 14 and 15, Fig. 3*a*). The protein was mixed with the DNA (50 µg in 50 µl) and incubated at 30 °C for 10 min in 50 mM Tris-HCl pH 8, 100 mM NaCl, 2 mM EDTA buffer. The reaction product was sedimented through a 10–30% glycerol gradient in the same buffer. Centrifugation was in a SW41 Ti rotor at 30,000 r.p.m. for 19 h at 4 °C. 500 µl fractions were collected and 10 µl per fraction was used to measure the radioactivity. The DNA content of each fraction was determined by measuring the absorbance at 260 nm. The control gradient contained only the protein and no DNA. *b*, To analyse the interaction of Pr76^{gag} with DNA, the same procedure was applied. Pr76^{gag} and other p19-related proteins isolated by immune affinity chromatography (see Fig. 3*b*) were dialysed and sedimented in the same conditions as the 110K protein in the absence of DNA. *c*, Filter binding assay of 110K and Pr76^{gag} proteins and DNA. The ability of the purified 110K and Pr76^{gag} proteins (fraction 14, Fig. 3*a* and *b*) to bind to DNA were analysed using a filter binding assay¹⁶. The incubation mixture (0.5 ml) contained 50 mM Tris-HCl pH 8.0, 2 mM EDTA, 2 µg of sheared ³H-labelled double-stranded DNA from chicken embryo fibroblasts, corresponding to 155,000 c.p.m., and proteins as indicated. After 10 min at 30 °C, the mixture was slowly filtered through a nitrocellulose filter (Selectron, Schleicher & Schuell, type BA85) and washed with 10 ml of the buffer used in the incubation mixture. The filter was pretreated with 0.3 M NaOH for 5 min and washed with H₂O (ref. 22). Binding of the proteins to the filter was controlled by determination of the ³⁵S-methionine radioactivity bound to the filters (top).



the filter (Fig. 4*c*, top). From this assay it was estimated that ~1 µg of 110K protein bound to 1 µg of cellular DNA, based on specific activities of 5×10^6 c.p.m. per mg of cellular proteins and 8×10^4 c.p.m. per µg of DNA.

To rule out the possibility that binding of the 110K protein to DNA was a consequence of the stringent elution conditions from the immune affinity column and not a property of the native protein, DNA binding to 110K was analysed by a modified approach. Instead of eluting the protein from the immune affinity column described in Fig. 3, the loaded column was exposed to ³H-labelled cellular DNA. In these conditions 85% of the DNA was bound to the column and did not elute with 0.6 or 2 M NaCl; only treatment with 1% SDS allowed its elution (Fig. 5*a*). A subsequent treatment of the column with pH 2 buffer removed the ³⁵S-methionine-labelled 110K protein. When the column was then analysed in a control experiment for its binding of DNA without the 110K protein, most of the DNA did not bind in these conditions (Fig. 5*b*).

We describe here for the first time the isolation and characterization of a transformation-specific protein of an acute avian leukaemia virus. It differs from other known oncornaviral proteins¹¹ in its nuclear localization and ability to bind to DNA. None of the gag-related polyproteins shown in Fig. 3*b* bound to DNA, indicating that DNA binding is probably due to the

transforming part. However, the biological significance of this property is unknown. Furthermore, transformation of fibroblasts analysed here will have to be compared with that of bone marrow cells, the *in vivo* target cells which differ in their transformation parameters¹⁷. The properties of 110K protein—although novel for oncornaviral transformation-specific proteins—closely resemble those of the large T antigen of SV40¹⁸.

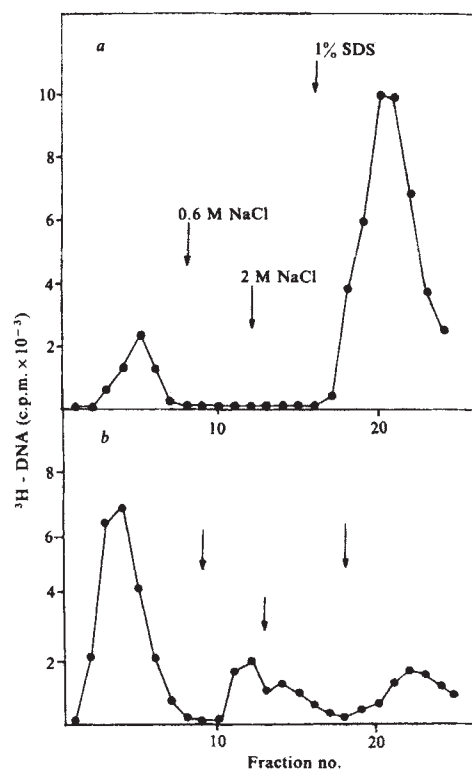


Fig. 5 Interaction of DNA and immune affinity column-bound 110K protein. *a*, An immune affinity column was prepared and loaded with ³⁵S-methionine-labelled cellular lysate as described in Fig. 3 legend. Instead of eluting the 110K protein from the column, ³H-labelled cellular DNA (7.5×10^5 c.p.m., specific activity 10^5 c.p.m. per µg) was loaded onto the column. A wash with PBS removed 15% of the DNA from the column. It was then treated with 0.5 ml TE buffer (50 mM Tris-HCl pH 8.0, 2 mM EDTA) containing 0.6 M NaCl and 2 M NaCl, respectively. No DNA was eluted. Only with TE buffer containing 1% SDS, 85% of the total DNA input radioactivity was recovered. 0.1-ml fractions were collected, 5 µl of each fraction being used to determine the radioactivity. The column was then treated with pH 2 buffer to elute the 110K protein as described in Fig. 3*a* legend. Subsequently the column was rinsed with PBS and used for a control experiment (*b*). The same amount of ³H-labelled DNA was applied to the column and treated exactly as described in *a*. 65% of the input DNA appeared in the flow-through, 25% eluted with 0.6 M and 2 M NaCl and residual 10% with 1% SDS.

Other gag-onc fusion proteins can be isolated in the way described here and their properties compared. Also, the availability of the purified 110K protein will permit polyvalent or monoclonal antibodies to be raised—preferably after removal of the gag portion by p15^{19,20}—so that transformation-specific antibodies will become available.

We thank M. K. Owada, T. Bunte and Erich Lanka for stimulating discussions and cooperation, and Klaus Bister for supplying the MC29-Q8-NP cells.

Received 5 November 1981; accepted 29 January 1982.

- Bishop, J. M. *Cell* **23**, 5–6 (1981).
- Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J. M. & Astrin, S. M. *Cell* **23**, 323–334 (1981).
- Payne, G. S. *Cell* **23**, 311–322 (1981).
- Hayward, W. S., Neal, B. G. & Astrin, S. M. *Nature* **290**, 475–480 (1981).
- Bister, K., Hayman, M. J. & Vogt, P. K. *Virology* **82**, 431–448 (1977).
- Kitchener, G. & Hayman, M. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1637–1641 (1980).
- Mellon, P., Pawson, A., Bister, K., Martin, G. S. & Duesberg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5874–5878 (1978).
- Greiser-Wilke, I., Owada, M. K. & Moelling, K. *J. Virol.* **39**, 225–329 (1981).
- Johnson, G. D., Holborow, E. J. & Darling, J. in *Handbook of Experimental Immunology* Vol. 1 (ed. Weir, D. M.) Ch. 15 (Blackwell, Oxford, 1978).
- Friis, R. R. *Virology* **50**, 701–712 (1972).
- Hynes, R. O. *Cell* **21**, 601–602 (1980).
- Penman, S. in *Fundamental Techniques in Virology* (eds Habel, K. & Salzman, N. P.) 35–48 (Academic, New York, London, 1969).
- Gersten, D. M. & Marchalonis, J. J. *J. immun. Meth.* **24**, 305–309 (1978).
- Secher, D. S. & Burke, D. C. *Nature* **285**, 446–450 (1980).
- Marmur, J. *J. molec. Biol.* **3**, 208–218 (1961).
- Davis, J., Sherer, M., Tsai, W. P. & Long, C. J. *J. Virol.* **18**, 709–718 (1976).
- Ramsey, G., Graf, T. & Hayman, M. J. *Nature* **288**, 170–172 (1980).
- Myers, R. M., Rio, D. C., Robbins, A. K. & Tijian, R. *Cell* **25**, 373–384 (1981).
- Dittmar, K. E. J. & Moelling, K. *J. Virol.* **28**, 106–118 (1978).
- Moelling, K., Scott, A., Dittmar, K. E. J. & Owada, M. K. *J. Virol.* **33**, 680–688 (1980).
- Kessler, S. W. *J. Immun.* **115**, 1617–1624 (1975).
- Otto, B., Baynes, M. & Knippers, R. *Eur. J. Biochem.* **73**, 17–24 (1977).

Spontaneous fragmentation of actin filaments in physiological conditions

Albrecht Wegner

Department of Physiological Chemistry I, Ruhr-University Bochum, Universitätsstrasse 150, D-4630 Bochum, FRG

Actin filaments have been shown to result from nucleation and consecutive binding of actin monomers to the ends of filaments¹. Nucleation is assumed to consist of the aggregation of a few monomers to form a small filament, but apart from the nucleation and the elongation reaction, end-to-end association of filaments has also been reported². It is not known whether in physiological conditions actin filaments can break spontaneously at subunit contacts along the filament. In view of the possible importance of spontaneous fragmentation in the determination of the length, number and turnover of thin filaments in the cell, I have now studied this reaction by measuring the polymerization kinetics of actin at low total concentrations in physiological conditions (1 mM MgCl₂, 100 mM KCl, 37 °C). Quantitative analysis of the polymerization curves suggests that actin filaments break spontaneously and consequently create more nucleation sites as the reaction proceeds. This explains the autocatalytic nature of the polymerization kinetics.

The polymerization was followed by 90° light-scattering intensity using a Zeiss fluorometer (546 nm). The light-scattering intensity R of long rod-like aggregates such as actin filaments has been shown to be proportional to the concentration of actin filament subunits c_f (refs 3,4).

$$R = a \cdot c_f \quad (1)$$

The constant a was calibrated by measuring the scattering intensity of solutions of known actin filament subunit concentrations. Actin was prepared as described earlier⁴, thereby omitting

the modification with *N*-ethylmaleimide, and the protein was applied to a Bio-Gel P-150 column to remove aggregates. Actin (<10 μM) was dialysed against a buffer containing 0.5 mM ATP, 10 μM MgCl₂, 1 mM dithiothreitol, 200 mg l⁻¹ NaN₃ and 5 mM triethanolamine hydrochloric acid (pH 7.5). The polymerization was initiated by mixing two volumes of dialysed monomeric actin with one volume of a buffer consisting of 0.5 mM ATP, 1.5–6 mM MgCl₂, 150–600 mM KCl, 0–150 μM CaCl₂, 0–600 μM EGTA, 10 mM dithiothreitol, 200 mg l⁻¹ NaN₃ and 5 mM triethanolamine hydrochloric acid (pH 7.5). Before mixing, all solutions were centrifuged for 1 h at 100,000g to remove dust and possibly occurring actin polymers. All experiments were performed at 37.1 °C.

The polymerization curves are depicted in Fig. 1. An attempt was made to fit these curves by a simple polymerization model⁴ that takes into account nucleation (K_N , equilibrium constant for formation of a nucleus from n monomers), association of monomers to the filament ends (k , association rate constant) and dissociation of subunits from the filament ends (k' , dissociation rate constant). However, it was not possible to find calculated curves that were in good agreement with measured kinetics (Fig. 1). Particularly at low total actin concentrations, the measured curves are significantly more sigmoidal than the calculated curves.

The long lag phase observed at low total actin concentrations is indicative of an autocatalytic reaction such as self-reproduction of filaments by fragmentation. To test whether the experimental data are compatible with fragmentation of filaments, a kinetic theory can be readily developed by extension of the nucleation elongation model⁴. Filament subunits (concentration c_f) result from binding of monomers (concentration c_1) to the ends of filaments and are consumed by dissociation of subunits from the ends. Hence

$$\frac{dc_f}{dt} = (kc_1 - k')C = (k(c_{\text{tot}} - c_f) - k')C \quad (2)$$

where c_{tot} is the total actin concentration and C is the concentration of actin filaments. The rate of filament formation by

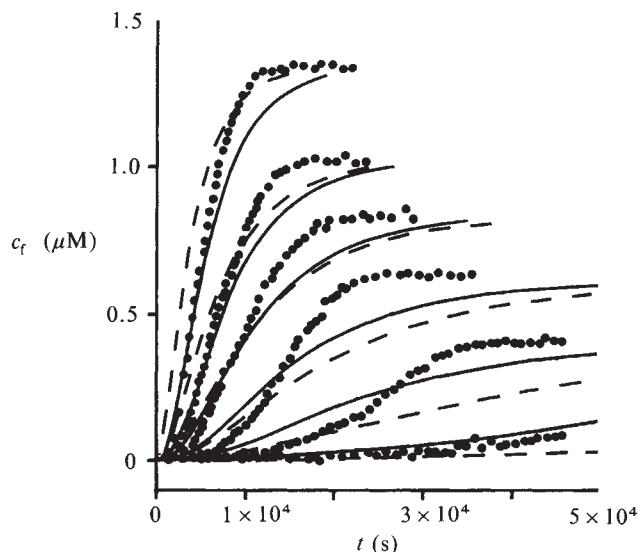


Fig. 1 Time course of the actin filament subunit concentration c_f . Salt concentrations: 1 mM MgCl₂ and 100 mM KCl. ●, Measured kinetics; total actin concentrations (μM): 0.41, 0.62, 0.82, 1.03, 1.23 and 1.54; final constant monomer concentration (critical concentration): $\bar{c}_1 = k'/k = 0.21$ μM. —, Calculated curves; the nucleus was assumed to be a dimer ($n=2$) and fragmentation was assumed not to occur; rate parameter obtained by combining equations (2) and (3): $k'^2 K_N = 1 \times 10^3 \text{ mol}^{-1} \text{ s}^{-2}$. ---, Calculated curves; the nucleus was treated as a tetramer ($n=4$) and fragmentation was assumed not to occur; rate parameter: $k'^2 K_N = 1.1 \times 10^9 \text{ mol}^{-3} \text{ s}^{-2}$.

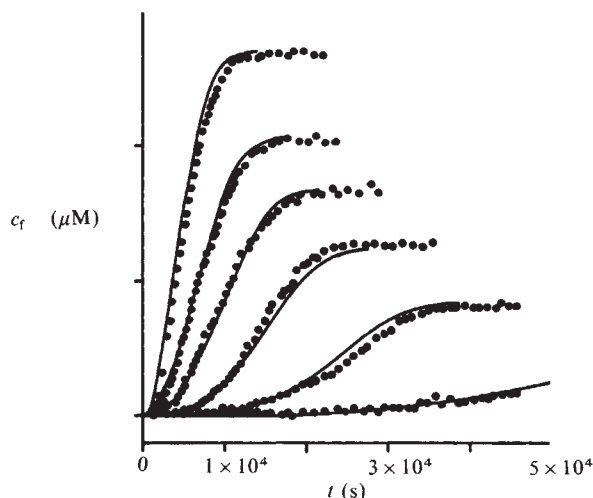


Fig. 2 Time course of the actin filament subunit concentration c_f . For salt concentrations and actin concentrations see Fig. 1 legend. ●, Measured curves; —, Calculated curves; the nucleus was treated as a tetramer ($n = 4$) and fragmentation was assumed to occur; rate parameters obtained by combining equations (2) and (3): $k^2 K_N = 5.8 \times 10^8 \text{ mol}^{-3} \text{ s}^{-2}$, $k' k_{fr} = 1.7 \times 10^{-8} \text{ s}^{-2}$.

nucleation and fragmentation is given by

$$\frac{dc_f}{dt} = (k(c_{tot} - c_f) - k')K_N(c_{tot} - c_f)^n + k_{fr}c_f \quad (3)$$

k_{fr} is the rate constant of fragmentation at a subunit contact along an actin filament. The first term of the kinetic equation (3) represents nucleation from monomers⁴ whereas the second term gives the rate of filament formation by fragmentation. Because of the uniformly double-helical structure of actin filaments the fragmentation reaction is assumed to occur at any subunit contact with the same rate, thereby neglecting end effects. When fragmentation is assumed to occur, polymerization curves can be calculated that reveal a long lag phase at low total actin concentrations and are in good agreement with the measured data (Fig. 2).

Slow isomerization steps in the early phases of polymerization can be excluded as a reason for the lag phase because isomerization steps lead to a lag phase independent of the total actin concentration whereas the experimental curves have increased sigmoidicity only at low total actin concentrations.

Addition of calcium (50 μM) or EGTA (200 μM) or variations in the magnesium (0.5–2 mM) or potassium concentration (50–200 mM) did not significantly affect the shape of the polymerization curves. The experiments suggest that actin filaments can break spontaneously quite independently of the ion concentrations.

The inherent instability of actin filaments resulting in spontaneous fragmentation may provide a further pathway of filament formation in addition to nucleation from actin monomers and growth of filaments induced by other molecules. This reaction may be important in determining the length and the number of filaments in the cell. Recently, several molecules^{5,6} have been reported to cause the conversion of filaments into short fragments by binding to one end of the filaments. The question arises of whether these molecules act by preventing end-to-end association of spontaneously broken filaments or whether they disrupt actin filaments.

The C conformation is a low salt form of sodium DNA

N. J. Rhodes*, A. Mahendrasingam*, W. J. Pigram*, W. Fuller*, J. Brahm†, J. Vergne† & R. A. J. Warren‡

* Department of Physics, University of Keele, Staffordshire ST5 5BG, UK

† Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire, Université Paris VII, 2 Place Jusieu, Paris Ve, France

‡ Department of Microbiology, University of British Columbia, Vancouver, British Columbia, Canada VGT 1W5

Three well-defined conformations have been observed for naturally occurring DNAs in oriented fibres. These have been designated A, B and C and their detailed geometries in a variety of crystalline and semi-crystalline forms determined by X-ray fibre diffraction analysis^{1–6}. The C conformation is commonly observed as a low humidity form of the lithium salt³. However, when the counter-ion is Na^+ rather than Li^+ , previous X-ray fibre diffraction studies^{6,7} have suggested that the C form is at best a poorly favoured conformation. Arnott and co-workers^{6,8} reported that the C form is stable at relative humidities and salt contents which are both intermediate between those which favour the A and B conformations. Zimmerman and Pfeiffer⁹ observed C-like patterns from fibres of Na-DNA immersed in *t*-butanol/water mixtures. They emphasized that the C form should be visualized as a family of related structures and suggested that continuous smooth transitions occur between the various members of the C family and between the B and C forms. IR linear dichroism studies have shown that a C-like conformation can occur in oriented films of Na-DNA¹⁰ at very low salt contents. Transitions between the A and B forms of Na-DNA depend on the salt content of the fibre and its relative humidity¹¹. Here we describe the conditions for routinely observing the C form of Na-DNA. These are applicable to a wide variety of natural DNAs and to the synthetic polynucleotide poly(dA-dC)·poly(dG-dT).

The DNAs studied were calf thymus, *Clostridium perfringens*, herring sperm, pollock roe, salmon sperm, *Escherichia coli* (all from Sigma), SP 15 (from Professor M. Mandel), T2 phage, *Micrococcus lysodeikticus* (both from Miles), $\phi\text{W-14}^{11}$, and poly(dA-dC)·poly(dG-dT). The final stage in the preparation was by either ethanol precipitation from $\leq 0.1 \text{ M}$ NaCl or centrifugation at 50,000 r.p.m. from NaCl as low as 0.001 M. Fibres were drawn from either the precipitate or the concentrated gel and X-ray diffraction patterns obtained at relative humidities from 0 to 98%. The NaCl concentration in a fibre was increased by re-dissolving it in a small quantity of distilled water and then adding to it a small quantity of a standard NaCl solution. A new fibre was pulled from the resulting gel, without loss of DNA. When fibres were made from precipitated material, the piece used was weighed and the DNA in the specimen determined after allowing for water content. For specimens from centrifuged material similar estimates were made using measurements of A_{260} for the original solution and supernatant. Hence, it was possible to calculate the number of Na^+ and Cl^- added per DNA PO_4^- .

All the Na-DNAs studied could be induced to assume the C form and Fig. 1 shows examples of their diffraction patterns. Where the patterns were sufficiently well defined they were of the hexagonal rather than orthorhombic type, with helical parameters not significantly different from those previously reported³. Conditions were found for which all the DNAs studied assumed the B conformation and also, with the exception of DNA from T2 and SP15, the A conformation. C patterns from the various Na-DNAs were only observed from fibres

Received 13 November 1981; accepted 29 January 1982.

1. Oosawa, F. & Kasai, M. *J. molec. Biol.* **4**, 10–21 (1962).
2. Nakaoka, Y. & Kasai, M. *J. molec. Biol.* **44**, 319–332 (1969).
3. Casassa, E. F. *J. chem. Phys.* **23**, 596–597 (1955).
4. Wegner, A. & Engel, J. *Biophys. Chem.* **3**, 215–225 (1975).
5. Yin, H. L. & Stosel, T. P. *Nature* **281**, 583–586 (1979).
6. Hasegawa, T., Takahashi, T., Hayashi, H. & Hatano, S. *Biochemistry* **19**, 2677–2683 (1980).

containing very little NaCl and in an environment of low relative humidity.

For all DNAs which could assume the A form, fibres giving the C pattern at low humidity exhibited transitions first from the C form to the fully crystalline A form and then from A to the semi-crystalline B form as the relative humidity was increased. Generally, the greater the NaCl content of the fibre the lower the relative humidity at which the A pattern occurred. For the lowest amounts of NaCl, C was usually observed at relative humidities from 32 to 75% but as the NaCl content was increased the upper limit in the relative humidity at which C occurred was reduced until it reached a level when the C pattern was no longer observed. For poly(dA-dC)·poly(dG-dT) this occurred when more than 0.7 Cl⁻ ions per DNA PO₄⁻ had been added to the fibre, but for natural DNAs rather less NaCl was required—typically 0.5 Cl⁻ per PO₄⁻. Below 32% relative humidity the X-ray diffraction patterns became much less well defined, indicating a progressive collapse of the structure. As reported previously¹¹, the relative humidity at which the A to B transition occurred decreased with increasing NaCl content. Generally, the transitions C→A→B were fully reversible as the relative humidity was reduced. However, for poly(dA-dC)·poly(dG-dT), with the NaCl concentrations so far studied, the A→B transition was reversible but the C form was only regained if the fibre was re-wetted to form a gel and a new

fibre drawn. For DNAs which undergo the C→A→B transitions we have observed patterns which are mixtures of A and C and of A and B. DNA from T2 and SP15 underwent a transition from C to semi-crystalline B with increasing relative humidity.

These results establish the C form as a major conformational possibility for the DNA double helix in biological environments. In contrast to previous work, we have shown that there is no difficulty in inducing Na-DNA to assume the C conformation if both the water content and the level of NaCl in the fibre are sufficiently low. *In vivo*, DNA can be expected to occur predominantly either as the sodium salt or in a state where the Na⁺ ions are displaced by basic groups such as polyamines and lysine and arginine side chains in proteins. Situations can therefore be envisaged in which the conformation of an active region of DNA is controlled by changes in its hydration and the concentration of Na⁺ and Cl⁻ ions in its vicinity.

Earlier reports suggested that the C form of DNA was a variant of the B conformation, with ready transition between the two forms. While such a transition is observed for Li-DNA, in this study of Na-DNA the C→B transition passes through an intermediate A conformation except for those DNAs that do not assume the A form. In contrast to the earlier claim^{6,8} that at low hydration the NaCl content of fibres of Na-DNA which gave the C form was intermediate between those appropriate for observing the A and B forms, we find that C is favoured by NaCl contents lower than those which favour A. Our view of the A form as an intermediate between the C and B forms of Na-DNA is based on two distinct sets of observations. First, the C→A→B transitions occur as a function of relative humidity for fibres containing small amounts of NaCl. Second, the C→A→B transitions occur at a fixed relative humidity (say 57%) as the NaCl content of a fibre is raised.

A striking feature of this study is the wide range of natural DNAs for which the C form was observed, particularly for DNAs with a G+C content of 31–72%, in agreement with IR studies¹⁰. Perhaps of even greater significance, the C form was observed for three DNAs (ϕW-14, SP15, T2) in which a substantial fraction of the nucleotides were extensively modified^{12–14}. The observation of the C form for this very wide range of natural DNAs suggests it is a structure sufficiently favoured stereochemically for it not to be destabilized by substantial changes in either its charge distribution or the addition of bulky groups. We did not observe the A form for T2 DNA, in accord with earlier studies which suggested this was a consequence of glucosylation of a large fraction of the cytosine residues¹⁵. That report also suggested that at low humidity T2 DNA adopts a novel conformation, the T form. In view of our observation of the C form as a low humidity conformation of T2 DNA and because the helix pitch and overall intensity distribution reported for the T form are broadly similar to those which characterize C patterns, we believe that the T form is a close relative, if not a member, of the C family of structures.

The various synthetic polynucleotides exhibit much more variety in the conformations they assume than do natural DNAs⁸. We have therefore concentrated here on the occurrence of the C conformation for the sodium salt of naturally occurring DNAs. However, we have also studied the synthetic polynucleotide poly(dA-dC)·poly(dG-dT) as an example of DNA with an 'average' nucleotide composition but a highly regular base sequence. Although the C form is also observed for fibres of this DNA containing low amounts of salt at low relative humidities, these fibres differ from those of natural DNAs in the irreversibility observed in the C→A transition. This is reproducible, having been observed for a number of different fibres and also be re-wetting the same fibre. This unusual behaviour of poly(dA-dC)·poly(dG-dT) may be a consequence of its particular nucleotide sequence, resulting in significantly different base-stacking contributions to the conformational energy, or to different patterns of solvent interactions around the DNA.

This work was supported by SERC (to W.F.) and NATO (to W.F. and R.A.J.W.).

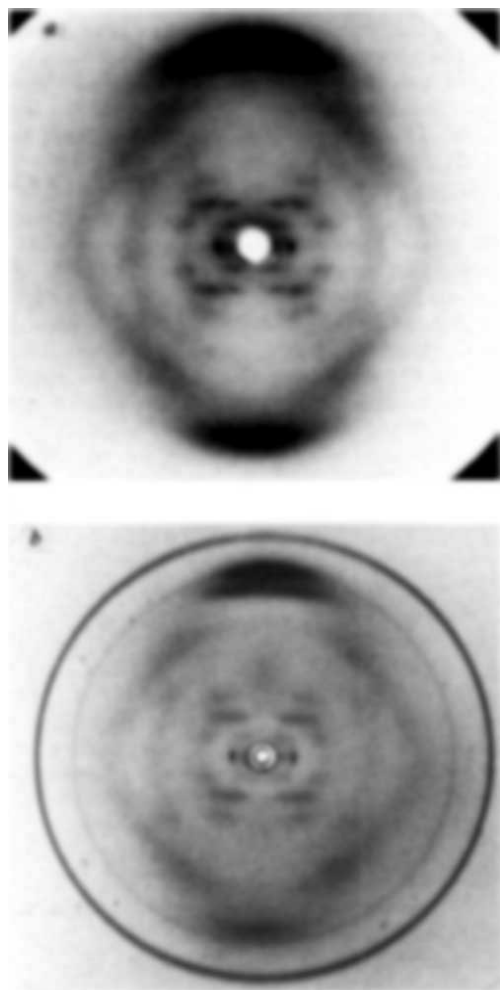


Fig. 1 X-ray diffraction patterns of the C type from calf thymus Na-DNA at 57% relative humidity (a) and Na poly(dA-dC)·poly(dG-dT) at 57% relative humidity (b) (the diffraction ring at 2.81 Å is from NaCl crystallites in the fibre). The two features generally regarded as characteristic of C patterns³ are a peak at $\sim 0.1 \text{ Å}^{-1}$ on the first layer line and strong diffraction close to the meridian on the eighth layer line. The peaks on the eighth and ninth layer lines in b are particularly well resolved.

Received 16 November 1981; accepted 28 January 1982.

- Fuller, W., Wilkins, M. H. F., Wilson, H. R. & Hamilton, L. D. *J. molec. Biol.* **12**, 60–80 (1965).
- Langridge, R. *et al. J. molec. Biol.* **2**, 38–64 (1960b).
- Marvin, D. A., Spencer, M., Wilkins, M. H. F. & Hamilton, L. D. *J. molec. Biol.* **3**, 547–565 (1961).
- Arnott, S. & Hukins, D. W. L. *Biochem. biophys. Res. Commun.* **47**, 1504–1509 (1972).
- Arnott, S. & Hukins, D. W. L. *J. molec. Biol.* **81**, 93–105 (1973).
- Arnott, S. & Selsing, E. *J. molec. Biol.* **98**, 265–269 (1975).
- Bram, S. & Baudy, P. *Nature* **250**, 414–416 (1974).
- Leslie, A. G. W., Arnott, S., Chandrasekaran, R. & Ratliff, R. L. *J. molec. Biol.* **143**, 49–72 (1980).
- Zimmerman, S. B. & Pfeiffer, B. H. *J. molec. Biol.* **142**, 315–330 (1980).
- Brahms, J., Pilet, J., Tan, T. P. L. & Hill, L. R. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3352–3355 (1973).
- Cooper, P. J. & Hamilton, L. D. *J. molec. Biol.* **16**, 562–563 (1966).
- Kropinski, A. M. B., Bose, R. J. & Warren, R. A. J. *Biochemistry* **12**, 151–157 (1973).
- Warren, R. A. J. *A. Rev. Microbiol.* **34**, 137–157 (1980).
- Wyatt, G. R. & Cohen, S. S. *Biochem. J.* **55**, 774–782 (1953).
- Mokulskii, M. A., Kapitonova, K. A. & Mokul'skaya, T. D. *Molec. Biol.* **6**, 716–731 (1972).

Transfection of cells from a xeroderma pigmentosum patient with normal human DNA confers UV resistance

Toshiya Takano, Makoto Noda* & Taka-aki Tamura

Department of Microbiology, Keio University School of Medicine, Shinjuku-Ku, Tokyo-160, Japan

Xeroderma pigmentosum (XP) is a human genetic disorder characterized clinically by extreme photosensitivity of the skin and eyes and a high incidence of skin cancer. Cell lines derived from XP patients are very sensitive to UV irradiation, and it is thought that the biochemical lesion responsible for this sensitivity is a defect in excision repair of UV-induced DNA damage¹. We report here that transfection of cells from an XP patient with DNA from normal human cells can produce cells resistant to UV damage. The resistant cells have presumably acquired a gene conferring UV resistance; this study represents the first step towards the isolation and characterization of the gene.

Skin fibroblasts of a XP patient, XP20S, were transformed by simian virus 40 (SV40) and established as a permanent cell line, XP20S(SV), by Takebe *et al.*². XP20S belongs to the complementation group A and shows an extremely low level of unscheduled DNA synthesis after UV irradiation². XP20S(SV), which has the characteristics of a cultured XP cell as well as those of the parental skin fibroblasts (it was highly UV-sensitive), was used as the transfection recipient.

High-molecular weight DNA was extracted and purified from human cultured cells or from *Escherichia coli* K-12 by sodium lauryl sulphate–phenol treatment after digestion with pronase E and ribonuclease (ref. 3); 10–30 µg of each DNA were transfected to 7×10^5 cells of XP20S(SV) using a modification of the calcium phosphate precipitation method of Wigler *et al.*⁴. The transfected cultures were then irradiated with UV light and the UV-resistant (UV^r) cells were scored by colony formation. The transfection and selection of UV^r cells are described in Table 1 and the results are shown in Table 1a. No colonies were detected after mock transfection ('control' in Table 1a), or after transfection with the self DNA of XP20S(SV) or with *E. coli* DNA. UV^r colonies were detected in the cultures transfected with DNAs from human cell lines of rhabdomyosarcoma, A204 (ref. 5), and embryonic lung fibroblasts, MRC-5 (ref. 6), the response being relatively dose dependent. The cells isolated from 10 individual colonies per dish showed UV-resistance equal to that of the DNA donors, A204 and MRC 5. Figure 1 shows the UV sensitivity of the transformed XP20S(SV) cells. However, note that 3–6% of the colonies were very small and consisted of <10 individual cells. The cells from these minute colonies did not survive after colony isolation, and seemed to

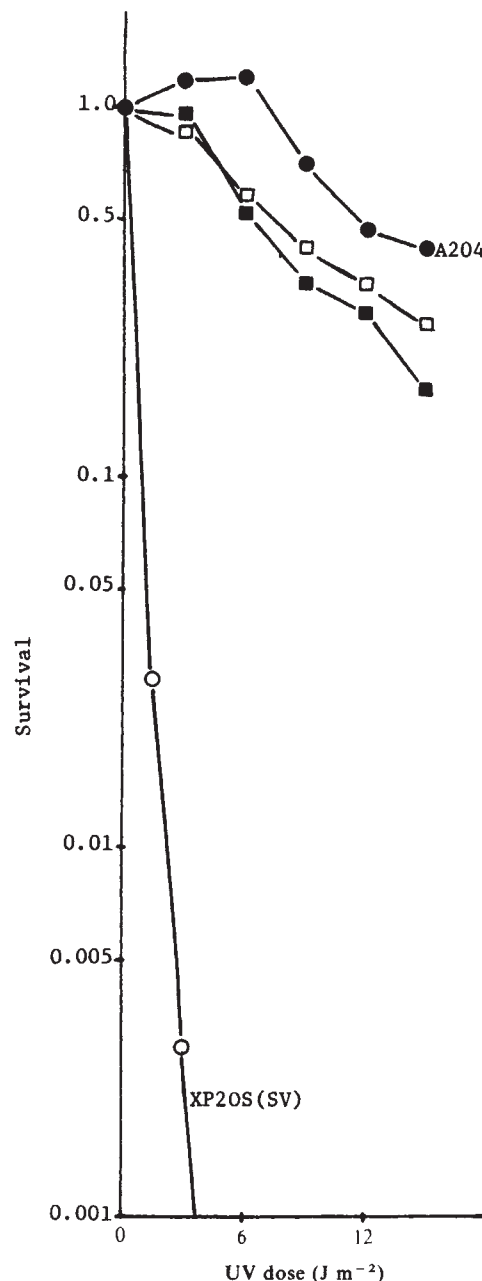


Fig. 1 UV-sensitivity of the transformed XP20S(SV) clones □, XP UV^r clone 1; ■, XP UV^r clone 2, and the parental ○, XP20S(SV) DNA donor cells (●, A204). In the experiment shown in Table 1a, colonies were isolated and 2×10^5 cells per colony were plated out in a 60-mm Petri dish, then irradiated with various doses of UV light from the bottom of the dish. The irradiated cells were cultured in the growth medium at 37 °C for 5 days, then the numbers of surviving cells were determined after trypsinization and Trypan blue staining. Results shown for XP UV^r clones 1 and 2 are typical of the cells of two representative colonies from the culture transfected with A204 DNA (20 µg per dish).

be practically UV-sensitive as were most of the recipient cells. Such minute colonies were detected neither in the mock infection nor in the dishes transfected by self and *E. coli* DNAs, which means that all the UV-sensitive XP cells were killed and UV^r transformants alone were detected in the experimental conditions used. The activity transforming the XP cell to UV^r was carried on the DNA molecules introduced into the transformed cells, because after treatment of A204 DNA with DNase I, there were virtually no UV^r transformants (Table 1a).

The frequency of the transformed UV^r colonies was 1.6–4.1 per µg DNA, whereas Wigler *et al.*⁴ described a transformation

* Present address: Laboratory of Tumor Virus Genetics, National Cancer Institute, Bethesda, Maryland 20205, USA.

frequency of ~10 colonies per 20 µg of DNA for the thymidine kinase and adenine phosphoribosyltransferase genes. The higher transformation frequency in our study seemed to be due to cell growth during the 6-day incubation and to re-plating of the cultures before selecting the UV^r colonies. The numbers of cells in the transfected cultures increased ~4–5 times during the incubation and the generation time of XP20S(SV) was ~26 h in the culture conditions used. Therefore, the initial transformation frequency of XP20S(SV) to UV^r seems to agree with that obtained by Wigler *et al.*⁴. The UV-sensitive phenotype of XP20S(SV) was stable; we have never detected any UV^r colonies in control cultures (untreated) in the experimental conditions used. Together, these results demonstrate that the UV-resistance in transformants is derived from the donor DNA and is not due to reversion of the recipient cell.

It was of interest to determine whether XP20S(SV) could be transformed with DNA which had previously been digested with restriction endonucleases. This would be useful with regard to cloning the gene which renders the XP cell UV-resistant. We therefore analysed the sensitivity of the transforming activity of A204 DNA digested with various restriction enzymes recognizing different hexanucleotide sequences at their

Table 1 Transformation of XP20S(SV) cells to UV-resistance by DNA transfection

Donor DNA	Amount of transfected DNA per dish (µg)	DNA digestion before transfection	UV-resistant colonies per 7 × 10 ⁵ recipient cells
<i>a</i>			
A204	30	—	123
	20	—	75
	10	—	39
MRC-5	30	—	98
	20	—	80
	10	—	16
XP20S(SV)	30	—	0
<i>E. coli</i>	30	—	0
A204	30	DNase I*	0
Control	—	—	0
<i>b</i>			
A204	20	None	78
A204	20	<i>SalI</i>	97.2
A204	20	<i>HindIII</i>	0 (9.2) [†]
A204	20	<i>XhoI</i>	142
A204	20	<i>BamHI</i>	0 (58) [†]
A204	20	<i>EcoRI</i>	0
Control	—	—	0

The recipient cells, XP20S(SV), were plated at 7 × 10⁵ per 60-mm Petri dish. The next day, DNA from each donor cell was transfected by the procedure described elsewhere⁴. After 6 h incubation with the DNA-calcium phosphate precipitates, the cells were treated with 10% dimethyl sulphoxide in HBS⁴ for 30 min at room temperature. Then the cells were washed once with HBS and incubated in Dulbecco's modified Eagle's minimum essential medium (Gibco) supplemented with 10% newborn calf serum (Mitsubishi Kasei Industrials) at 37 °C for 6 days. The culture medium was changed once a day. After incubation, the cultures were each trypsinized and subdivided into four 60-mm Petri dishes (Lux, no. 5213). The next day, the cells were rinsed twice with HBS and dried by keeping the dishes up-side down until the cells showed hardly any damage that could be observed during the subsequent incubation. They were then irradiated with UV light from both the top and bottom of the dishes, at 2.3 and 2.8 J m⁻², respectively. The irradiated cells were incubated in the growth medium at 37 °C for a further 6 days then the cultures were again irradiated with UV light at 1.5 J m⁻² from both sides of the dishes, by the procedure described above. After incubation in the growth medium for 2 weeks surviving colonies were scored after formaldehyde fixation and staining with Giemsa. The values shown are the averages of results from five independent experiments. *b*, A204 DNA was digested with restriction enzymes (Takara Shuzo or New England BioLabs); 2 units of enzyme per µg DNA, at 37 °C for 4 h. Complete digestion of DNA samples with these enzymes was monitored by adding a small amount of bacteriophage λ DNA to an aliquot of each reaction mixture, and analysing the digestion products by gel electrophoresis. The cleaved DNA samples were precipitated with 70% ethanol, dissolved in a small amount of 10 mM Tris-HCl, 1 mM EDTA pH 7.2, and transfected as described above.

* The DNA was treated with 1 mg ml⁻¹ pancreatic DNase I (Sigma, type I) in 10 mM Tris-HCl, 5 mM MgCl₂ pH 7.2 at 37 °C for 60 min and then transfected.

[†] Number of minute colonies detected (see also Fig. 2) in which cells did not survive after colony isolation.

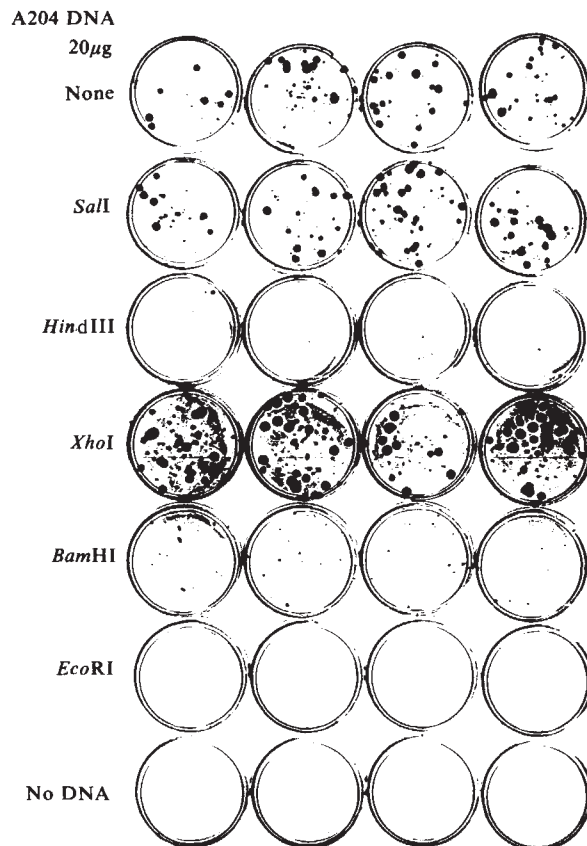


Fig. 2 Petri dishes containing the UV^r colonies detected in the experiments shown in Table 1*b*. The four dishes in each row were derived from one transfected culture. The restriction enzymes used to digest the transfecting DNA are indicated at the left-hand side.

cleavage sites (see Table 1*b* and Fig. 2). *SalI* and *XhoI* raised the transformation efficiency whereas *HindIII*, *BamHI* and *EcoRI* removed the transforming activity. On the other hand, considerable numbers of minute colonies were detected in the cultures treated with *BamHI*- and *HindIII*-digested DNAs; these digested DNAs seemed to raise slightly the UV-resistance of the recipient cell, possibly by abortive transformation. However, we were unable to characterize further the cells in these minute colonies because they did not survive after colony isolation.

The UV-resistant XP20S(SV) cells transformed by both high-molecular weight and restriction endonuclease-cleaved DNA from A204 and MRC-5, showed stable UV^r phenotypes during passage in culture; we found no revertants that were highly sensitive to UV irradiation.

These results suggest that the DNAs of non-XP human cells contain a gene(s) which confers the UV^r phenotype on the XP cell, and which carries the *BamHI*, *EcoRI* and *HindIII* sites. *SalI* and *XhoI* cleave only outside the functionally important regions of the gene. The gene is not necessarily the wild-type counterpart of the defective gene in the XP cell, but it could be any gene which can suppress the UV-sensitive phenotype of the XP cell. We are now cloning the DNA sequences carrying the transforming activity using the selection of UV^r transformants described in this study.

We thank Dr Hiraku Takebe for supplying XP20S(SV) cells. This work was supported by grants-in-aid from the Japanese Ministry of Education, Science and Culture.

Received 6 January; accepted 3 February 1982.

- Arlett, C. F. & Lehmann, A. R. *A. Rev. Genet.* **12**, 95–115 (1978).
- Takebe, H., Nii, S., Iida, M. & Utsumi, H. *Mutat. Res.* **25**, 383–390 (1974).
- Hughes, S. H. *et al. Cell* **15**, 1397–1410 (1978).
- Wigler, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1373–1376 (1979).
- Giard, D. J. *et al. J. natn. Cancer Inst.* **51**, 1417–1423 (1973).
- Jacobs, J. P., Jones, C. M. & Baille, J. P. *Nature* **227**, 168–170 (1970).

Regulation of DNA ligase activity by poly(ADP-ribose)

Debbie Creissen & Sydney Shall

Cell & Molecular Biology Laboratory, University of Sussex, Brighton, Sussex BN1 9QG, UK

Nuclear ADP-ribosyl transferase (ADPRT)^{1,2} catalyses the transfer of ADP-ribose from NAD⁺ to chromatin proteins to form mono-, oligo- and poly(ADP-ribose)-modified chromatin proteins³⁻⁷. The enzyme is entirely dependent on DNA for activity⁸ and is activated by nicks in the DNA⁹⁻¹¹. Both radiation and alkylating agents deplete cells of NAD which is used to make (ADP-ribose)_n (refs 1, 2, 12, 15-17), and this depletion is prevented by inhibitors of ADPRT^{13,15,18-25}. ADPRT inhibitors also retard DNA strand-rejoining after damage^{1,2,26,27}, and potentiate the lethality of DNA-damaging agents^{1,2,28}. *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine, an alkylating agent, transiently increases intracellular (ADP-ribose)_n over 100-fold¹⁷. These data suggest that (ADP-ribose)_n biosynthesis is necessary for effective repair of some types of DNA damage¹. The question remains of which step in DNA repair requires ADPRT activity. The enzyme inhibitors do not prevent incision events^{1,2,27} nor do they inhibit repair replication²⁹⁻³¹. Here we describe experiments which support the hypothesis that ADP-ribosylation activates DNA ligase activity. We suggest that ADPRT has a general function in the control of the breaking and rejoining of DNA.

Mammalian cells possess two forms or species of DNA ligase, which have been chromatographically separated³²⁻³⁹. DNA ligase I seems to be involved mainly in replication, while DNA ligase II may be a repair enzyme^{32,33,37-44}. DNA ligase activity may be extracted from cells with 0.5 M salt and Triton X-100, the enzyme activity being reasonably reproducible. In both control cells and cells treated with the monofunctional alkylating agent dimethyl sulphate, the maximum amount of enzyme activity is extracted by 300 mM KCl. Doubling the salt concentration to 600 mM does not extract any further enzyme activity. The enzyme activity may be estimated by measuring the amount of ³²P-5'-phosphate in nicked DNA that becomes inaccessible and therefore resistant to alkaline phosphatase because the DNA ligase activity has converted the 5'-phosphate to a 3',5'-phosphodiester. We found that on exposure of mouse leukaemic L1210 cells to dimethyl sulphate, the total DNA ligase enzyme activity increased ~2.3-fold (Table 1).

This increase in enzyme activity was prevented by all classes of ADPRT inhibitor, indicating that it is a consequence of ADPRT activity (Table 1). The specificity of the ADPRT inhibitors is indicated by the inability of their non-inhibitory analogues, aminobenzoic acid and nicotinic acid, to prevent the enzyme stimulation (Table 1). Treatment of cells with ADPRT inhibitors does not inhibit the basal DNA ligase activity or the basal or stimulated ligase activity when present in the ligase assay (data not shown).

Again, the stimulating effect on DNA ligase activity of prior exposure of the cells to the alkylating agent is seen following purification of the enzyme by ammonium sulphate precipitation followed by successive chromatography on phosphocellulose and hydroxylapatite columns (Fig. 1). The total recovered enzyme activity of the DNA ligase peak eluting from the phosphocellulose column increased from 3.05 to 5.42 pmol min⁻¹. Hydroxylapatite chromatography of the DNA ligase peak from the phosphocellulose column shows that ligase II is the enzyme predominantly responsible for the increase in activity (Fig. 1). Whereas the ligase I activity eluting from the hydroxylapatite column increased from 1.90 to 2.40 pmol min⁻¹, that of DNA

ligase II increased from 1.27 to 6.20 pmol min⁻¹, a fivefold increase.

An antiserum has been described by Soderh  ll and Lindahl³³ that specifically inhibits DNA ligase I from calf, human, rabbit and mouse tissues but does not affect DNA ligase II from the same sources to a detectable extent. We found that this antiserum, at a 1:8 dilution, inhibited our DNA ligase I enzyme activity (see Fig. 1) by 95% but inhibited ligase II enzyme activity (see Fig. 1) by only 11%, which we attribute to contamination of the ligase II peak by some ligase I material. We conclude that our two ligase enzyme activities are identical to those described by Soderh  ll and Lindahl³³.

The fact that we used a high salt concentration means that the increase in ligase II activity is probably not due to differential extraction. This possibility is further excluded by the data in Fig. 1 which show that the specific enzyme activity (units of enzyme activity per unit protein) increases sevenfold on exposure to dimethyl sulphate.

We next investigated whether the blocking effect of these ADPRT inhibitors on dimethyl sulphate-stimulated total enzyme activity was specific for DNA ligase II activity. We treated two sets of cells with 300 μ M dimethyl sulphate, one set being simultaneously exposed to 5 mM 3-aminobenzamide, then extracted and partially purified the DNA ligase activities as described earlier. Again, dimethyl sulphate increased the DNA ligase II activity (Fig. 2a). However, this increase did not occur when 3-aminobenzamide was present (Fig. 2b).

One possible explanation of the above results is that DNA ligase II itself, a closely associated protein or some other

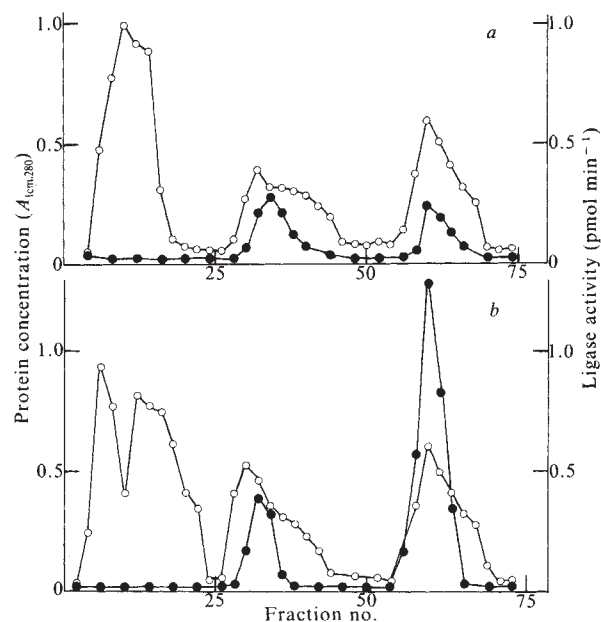


Fig. 1 Increase in DNA ligase II activity after exposure of L1210 cells to dimethyl sulphate. The cells (100 ml, 5×10^5 cells ml⁻¹, for each experiment) were incubated for 1 h with or without 300 μ M dimethyl sulphate. Salt extracts were prepared as described in Table 1 legend. Ammonium sulphate was added to 75% saturation. The precipitate was collected, redissolved in 15 mM potassium phosphate buffer pH 7.2 and dialysed against the same buffer overnight. A column of phosphocellulose (1.25 $\times$ 7 cm) pre-equilibrated with 15 mM potassium phosphate buffer pH 7.2 was loaded with the dialysed material and washed with 4 ml 15 mM potassium phosphate buffer pH 7.2. Elution with 800 mM NaCl, 15 mM potassium phosphate buffer pH 7.2 continued until no more protein was detected in the eluant. Fractions of 1.0 ml were collected and assayed for DNA ligase activity as described in Table 1 legend. The enzymatically active fractions from the phosphocellulose columns were pooled and loaded directly on to columns of hydroxylapatite (1.0 $\times$ 3 cm) pre-equilibrated with 15 mM potassium phosphate buffer pH 7.2. Stepwise elution was carried out with increasing concentrations of potassium phosphate buffer pH 7.2; 1 ml 15 mM, 4.0 ml 50 mM, 5.0 ml 140 mM, 5.0 ml 400 mM. Aliquots (200 μ l) were collected and assayed for protein concentration and DNA ligase enzyme activity. $\circ$, Protein concentration (A_{280} absorbance units); $\bullet$, DNA ligase activity (pmol min⁻¹). *a*, Cells not treated with dimethyl sulphate; *b*, cells exposed to 300 μ M dimethyl sulphate for 1 h.

Table 1 Stimulation of DNA ligase activity by an alkylating agent, and its prevention by inhibitors of ADPRT

Additions		Relative DNA ligase activity	
		Controls	Exposed to dimethyl sulphate
None	(10)	1.00	2.35
Benzamide	(2)	1.08	0.81
3-Aminobenzamide	(10)	0.80	1.08
5-Methylnicotinamide	(2)	1.10	0.85
Theophylline	(2)	0.94	1.06
Thymidine	(2)	0.94	0.97
3-Aminobenzoic acid	(2)	0.90	1.82
Nicotinic acid	(2)	1.06	1.82

Mouse lymphoblastic leukaemia cells (L1210 cells) (5×10^5 cells ml^{-1}) were treated with 300 μM dimethyl sulphate for 1 h in the presence or absence of the inhibitors (5 mM) or their analogues (5 mM). DNA ligase activity was then extracted from the cells. In control experiments the cells were exposed to the inhibitors or analogues, but were not treated with dimethyl sulphate. The cells were centrifuged and then washed twice with phosphate-buffered saline. The cell pellet was then resuspended in 1/10 the original volume of extraction buffer at 0°C. The extraction buffer consisted of 500 mM KCl, 20 mM Tris-HCl pH 7.5, 0.5% (v/v) Triton X-100. The cell pellet in the extraction buffer was sonicated twice in an MSE 150-W sonicator for 15 s each time. The enzyme extract was recovered by centrifugation at maximum speed in an Eppendorf microcentrifuge. The supernatant contained the DNA ligase activity; this supernatant was dialysed overnight against 20 mM Tris-HCl pH 7.5, 1 mM mercaptoethanol. DNA ligase activity was estimated with ^{32}P -5'-phosphate DNA as substrate, prepared according to ref. 45. The DNA ligase assay measures the rate at which ^{32}P -5'-phosphate becomes inaccessible to alkaline phosphatase because of DNA ligation. The enzyme assay mixture (200 μl) contained 75 mM Tris-HCl pH 7.5, 15 mM MgCl_2 , 200 μM ATP, 2 mM dithiothreitol, 20 μg bovine serum albumin, 20 pmol ^{32}P -5'-phosphoryl DNA (1,000 c.p.m. pmol^{-1}), 100 μg ml^{-1} protein from the enzyme extract. The assay mixture was incubated at 30°C for 30 min. The reactions were stopped by the addition of 5 μl EDTA and 15 μl 1.0 M Tris-HCl pH 8.5 at 0°C. The ^{32}P -phosphate that had become resistant to alkaline phosphatase was estimated according to the procedure described by Weiss *et al.*⁴⁶. One unit of DNA ligase activity catalyses the conversion of 1 pmol of ^{32}P -5'-phosphate to phosphatase-resistant material in 1 min in the above conditions. The DNA ligase assay is linear with time at least up to 60 min, regardless of whether the cell extracts have been exposed to dimethyl sulphate. Values in parentheses indicate the number of times each experiment was performed. All enzyme activities are expressed relative to the activity in cells not exposed to dimethyl sulphate, inhibitors or their analogues.

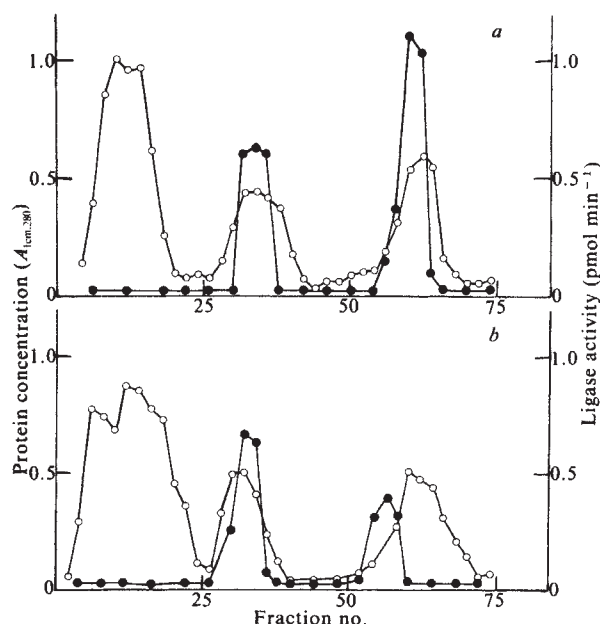


Fig. 2 3-Aminobenzamide prevents the increase in DNA ligase II activity observed after exposure of L1210 cells to dimethyl sulphate. The cells (100 ml, 5×10^5 cells ml^{-1} , for each experiment) were incubated for 1 h with 300 μM DMS. One set of cells was treated simultaneously with 5 mM 3-aminobenzamide. Salt extracts were prepared as described in Table 1 legend and partial purification of DNA ligases I and II was carried out exactly as described in Fig. 1 legend. ○, Protein concentration (A_{280} absorbance units); ●, DNA ligase activity (pmol min^{-1}). a, DNA ligase from cells exposed to 300 μM DMS; b, DNA ligase from cells exposed simultaneously to 300 μM DMS and 5 mM 3-aminobenzamide.

DNA-binding protein is ADP-ribosylated. Preliminary experiments indicate that when permeabilized cells are incubated with ^3H -NAD, a specific precursor of (ADP-ribose)_n, a peak of radioactivity co-migrates with the DNA ligase II activity on a hydroxylapatite column.

Thus, we observe that alkylation damage increases DNA ligase activity, predominantly that of DNA ligase II, the presumed repair enzyme. This increase in activity is totally prevented by inhibiting (ADP-ribose)_n biosynthesis.

(ADP-ribose)_n biosynthesis is required for effective DNA excision repair of some types of DNA damage¹, while ADP-ribose is required neither for the incision events^{1,2,28} nor for repair replication³⁰⁻³². We propose that (ADP-ribose)_n biosynthesis is required for effective DNA excision repair because it is necessary for efficient ligation. The nuclear ADPRT enzyme is stimulated by breaks in the DNA and subsequently stimulates their ligation. However, we suggest that (ADP-ribose)_n may have a more general role than that of DNA repair, in that it may be sensitive to breaks and regulate subsequent DNA ligation in DNA recombination, sister chromatid exchanges, gene rearrangements, transpositions and cell differentiation which may involve such changes in DNA structure.

This work was supported by the MRC, the SERC and the Cancer Research Campaign. We thank all our laboratory colleagues for constructive help and advice, and Dr S. Soderhäll and T. Lindahl for giving us their specific antiserum against DNA ligase I.

Received 5 October 1981; accepted 27 January 1982.

- Durkacz, B. W., Omidiji, O., Gray, D. A. & Shall, S. *Nature* **283**, 593-596 (1980).
- Durkacz, B. W., Nduka, N., Omidiji, O., Shall, S. & Zia'ee, A.-A. in *Novel ADP-ribosylations of Regulatory Enzymes and Proteins* (eds Smulson, M. E. & Sugimura, T.) 207-216 (Elsevier, Amsterdam, 1980).
- Sugimura, T. *Prog. Nucleic Acid Res. molec. Biol.* **13**, 127-151 (1973).
- Hiltz, H. & Stone, P. R. *Rev. Physiol. Biochem. Pharmac.* **76**, 1-58 (1976).
- Hayashi, O. & Ueda, K. A. *Rev. Biochem.* **46**, 95-116 (1977).
- Purnell, M. R., Stone, P. R. & Whish, W. J. D. *Biochem. Soc. Trans.* **8**, 215-227 (1980).
- Smulson, M. E. & Sugimura, T. (eds) *Novel ADP-ribosylations of Regulatory Enzymes and Proteins* (Elsevier, Amsterdam, 1980).
- Tsopanakis, C., Leeson, E., Tsopanakis, A. & Shall, S. *Eur. J. Biochem.* **90**, 337-345 (1978).
- Halldorsson, H., Gray, D. A. & Shall, S. *FEBS Lett.* **85**, 349-352 (1978).
- Berger, N. A., Weber, G. & Kaichi, A. S. *Biochim. biophys. Acta* **519**, 87-104 (1978).
- Benjamin, R. & Gill, M. J. *biol. Chem.* **255**, 10493-10508 (1980).
- Whish, W. J. D., Davies, M. I. & Shall, S. *Biochem. biophys. Res. Commun.* **65**, 722-730 (1975).
- Davies, M. I., Shall, S. & Skidmore, C. J. *Biochem. Soc. Trans.* **5**, 949-950 (1978).
- Davies, M. I., Halldorsson, H., Nduka, N., Shall, S. & Skidmore, C. J. *Biochem. Soc. Trans.* **6**, 1056-1057 (1978).
- Skidmore, C. J. *et al. Eur. J. Biochem.* **101**, 135-142 (1979).
- Jacobson, M. K., Levi, V., Juarez-Salinas, H., Barton, R. A. & Jacobson, E. L. *Cancer Res.* **40**, 1797-1802 (1980).
- Juarez-Salinas, H., Sims, J. L. & Jacobson, M. K. *Nature* **282**, 740-741 (1979).
- Preiss, J., Schlaeger, R. & Hiltz, H. *FEBS Lett.* **19**, 244-246 (1971).
- Clark, J. B., Ferris, G. M. & Pinder, S. *Biochim. biophys. Acta* **238**, 82-85 (1971).
- Shall, S. *et al. Biochem. Soc. Symp.* **42**, 103-116 (1977).
- Claycomb, W. C. *Biochem. J.* **154**, 387-393 (1976).
- Levi, V., Jacobson, E. L. & Jacobson, M. K. *FEBS Lett.* **8**, 144-146 (1978).
- Khan, M. G. thesis, Univ. Sussex (1977).
- Shall, S. *J. Biochem., Tokyo* **77**, 2p (1975).
- Purnell, M. R. & Whish, W. J. D. *Biochem. J.* **185**, 775-777 (1980).
- Durkacz, B. W., Irwin, J. & Shall, S. *Eur. J. Biochem.* **121**, 65-69 (1981).
- Gray, D. A., Durkacz, B. W. & Shall, S. *FEBS Lett.* **131**, 173-177 (1981).
- Nduka, N., Skidmore, C. J. & Shall, S. *Eur. J. Biochem.* **105**, 525-530 (1980).
- Durkacz, B. W., Irwin, J. & Shall, S. *Biochem. biophys. Res. Commun.* **101**, 1433-1441 (1981).
- James, M. R. & Lehmann, A. R. in *DNA Repair, Chromosome Alterations and Chromatin Structure* (eds Natarajan, A. T., Altman, H. & Obe, G.) (Elsevier, Amsterdam, in the press).
- Miwa, M. *et al. Biochem. biophys. Res. Commun.* **100**, 463-470 (1981).
- Söderhäll, S. & Lindahl, T. *Biochem. biophys. Res. Commun.* **53**, 910-916 (1973).
- Söderhäll, S. & Lindahl, T. *J. biol. Chem.* **250**, 8438-8444 (1975).
- Teraoka, H., Shimoyachi, M. & Tsukada, K. *FEBS Lett.* **54**, 217-220 (1975).
- Evans, D. G., Ton, S. H. & Keir, H. M. *Biochem. Soc. Trans.* **3**, 1131-1132 (1975).
- Zimmerman, S. B. & Levin, C. J. *J. biol. Chem.* **250**, 149-155 (1975).
- Söderhäll, S. *Nature* **260**, 640-642 (1976).
- Söderhäll, S. & Lindahl, T. *FEBS Lett.* **67**, 1-8 (1976).
- Teraoka, H., Shimoyachi, M. & Tsukada, K. *J. Biochem.* **81**, 1253-1260 (1977).
- Inoue, N. & Kato, T. *Proc. Japan Acad.* **52**, 461-464 (1976).
- Nakaya, N., Sawasaki, Y., Teraoka, M., Nakajima, H. & Tsukada, K. *J. Biochem.* **81**, 1575-1577 (1977).
- Mezzina, M. & Nocentini, S. *Nucleic Acids Res.* **5**, 4317-4328 (1978).
- David, J. C., Vinson, D., Letresne, J. & Signoret, J. *J. Cell Differentiation* **8**, 451-459 (1979).
- David, J. C., Vinson, D. & Fedacka, Bruner, B. *Exp. Cell Res.* **130**, 137-146 (1980).
- Weiss, B., Live, T. R. & Richardson, C. C. *J. biol. Chem.* **243**, 4530-4542 (1968).
- Weiss, B., Jacquemin-Sablon, A., Live, T. R., Fareed, G. C. & Richardson, C. C. *J. biol. Chem.* **243**, 4543-4555 (1968).

BOOK REVIEWS

From China with love

J.D. Bjorken

IN THE spring of 1979 T.D. Lee visited China and within the span of seven weeks delivered a remarkably intense set of lectures on elementary particles and statistical mechanics. The number of lecture hours averaged about three per day, seven days a week, and the size of the audience exceeded 900. Lee reports that "... it was very lively. My vocal chords, however, had problems afterward". Parts of the lecture notes have been published in Chinese and the particle physics portion has evolved into this volume.

It is an especially propitious time for such a course to appear. Because of the recent progress in understanding the nature of the fundamental quark and lepton constituents of matter and of the forces they exert on each other, we now are entering what should be a golden age of pedagogy. A few good texts have appeared and more are on the way. The ideal one is yet to be written. But Lee's lectures, as one might anticipate, set a towering standard of excellence. Within 820 pages of lucid text lie many of the important topics in elementary particle theory.

The central element of the new pedagogy is the belief that the strong and weak forces are described by generalizations of quantum electrodynamics called non-abelian gauge theories. A good understanding of this subject requires mastery of a high level of theoretical formalism. This is not easily found in older texts. Lee's lectures help to bridge the gap between the old and the new, concentrating on the theory of the strong force, called quantum chromodynamics or QCD.

But physical insights have not been sacrificed for formalism. The main physical features of the QCD picture of the inter-quark force are that it is relatively weak at short distances, and that at large distances it confines quarks — that is, only certain colour-neutral combinations of quarks and their antiparticles can exist in isolation. Lee beautifully elucidates these concepts, albeit with considerable subjectivity on the confinement question. At present, the confinement problem is not closed. There are several approaches extant, such as lattice QCD, instantons and string models. Only one approach, which makes an analogy of the QCD vacuum state with a perfect dielectric medium, is developed in the book. This reflects Lee's current research interests, and is a characteristic of the lectures. They are selective, not encyclopaedic.

The weak — or, better, electroweak —

Particle Physics and Introduction to Field Theory. By T.D. Lee. Pp.865. ISBN hbk 3-7186-0032-3; ISBN pbk 3-7186-0033-1. (Harwood: 1981.) Hbk \$59.50, Dfl.170; pbk \$19.50, Dfl.60.

force is now also generally believed to be described by a non-abelian gauge theory. Lee's lectures emphasize the older phenomenological foundations, although the prevailing $SU(2) \otimes U(1)$ "standard model" is given a spare but reasonable and up-to-date treatment. Here the computational technology might have been carried further. The physics might have been as well: there is no discussion of the ideas of the grand unification of strong, weak and electromagnetic forces, a topic now inspiring a considerable amount of experimental and theoretical work.

An important link between experiment and the basic gauge-theory concepts is the so-called quark-parton model. This is a simple manifestation of the idea that point-like quark constituents reside within nucleons and other hadrons. The quark-parton model is also an element of the new

pedagogy — one which is simple to apply and of great usefulness. It is a good way to learn a lot of physics easily and quickly. Lee develops it well, with relevant examples worked through in physically transparent ways.

But the core of Lee's lectures lies not in the new pedagogy, but the old. There is a concise introduction to quantum field theory, followed by an extended, masterful and definitive section on symmetry principles and the associated conservation laws. This leads to a full discussion of the neutral kaon system — that glory of wave mechanics and source of so many of our deepest insights into nature.

This volume will serve superbly as a comprehensive course in elementary particle theory at the advanced graduate-school level. It is possible to differ with Lee on questions of balance and of selection of material. But the pages are filled with his unique insights, pedagogical clarity and thoroughness. The book is a treasure. □

J.D. Bjorken is at the Fermi National Accelerator Laboratory, Batavia, Illinois.

A cooperative venture in soil science

Harry Vine

Characterization of Soils in Relation to their Classification and Management for Crop Production: Examples from Some Areas of the Humid Tropics. Edited by D.J. Greenland. Pp.446. ISBN 0-19-854538-X. (Clarendon/Oxford University Press: 1981.) £35, \$74.

THIS volume is, with a few added contributions, the report of a "benchmark soils" project which was based at the International Institute for Tropical Agriculture (IITA) at Ibadan, Nigeria. The core of the book, then, consists of detailed studies of 38 Nigerian soil profiles, in which about 20 specialists in soil science took part.

A major aim of the project was "to provide data that might enable existing soil classification systems used within the tropics to be improved"; this related especially to the American system, which was developed through successive *Approximations* to the massive handbook, *Soil Taxonomy*, issued by the United States Department of Agriculture in 1975. The profiles are discussed from this point of

view by F.R. Moormann, who points out a number of difficulties in the published scheme for tropical soils, particularly with regard to weatherable sand-size minerals and to the relative importance of the composition of the clay fraction and evidence of its movement in the profile. A fundamental contribution to the characterization of such soils is made by A.J. Herbillon in a review in which he examines both materials and dynamics.

Underlying the work on this range of profiles was the need, in the Farming Systems Programme of IITA, to discern more definitely what measurable properties of humid tropical soils may be relevant to maintenance of fertility in intensified arable cropping, efficient use of fertilizers and avoidance of erosion. In their review chapter, B.T. Kang and R.L. Fox aptly refer to "misconceptions" that these soils "are quite infertile and have management problems which are encountered nowhere else". On phosphate fixation problems, greatly exaggerated in textbooks, good progress is reported by

P.H. Le Mare, who used isotopically-labelled phosphate in distinguishing strong bonding from exchange-adsorption in laboratory studies of the soils.

Tropical ecosystems have been described as "fragile", particularly in semi-arid regions. D.J. Greenland writes here and elsewhere about "fragile" tropical soils, using the term in a literal sense, implying that particular properties of the soils (not the intensity of the rainfall) were responsible for a tendency to high rates of erosion. Clearly almost all the bench-mark soils, except those of the amphibolite area, can have little stable structure in the topsoil, being sandy loams; but as the sand is medium to coarse the infiltration rate remains fairly high, and run-off and erosion need not be intense. In this connection, some years ago the IITA group devised a system of no-tillage, which involved mulching with plant material, applying herbicides and omitting the yam crop which is important locally. In advocating this they may have been too greatly influenced by the amount of erosion that occurred on experimental plots under rather absurd methods of cultivation; planting on ridges with cross-bars, as on older stations in Nigeria, would have been a better yardstick.

It is also unfortunate that despite the long-term prospects for cattle, with diseases being brought under control, animal husbandry and therefore alternation of arable crops with grass-legume mixtures — which do well in southern Nigeria — were not considered in the Farming Systems Programme. Consequently one of the most interesting features of this book is the inclusion of a long chapter by P.A. Sanchez, "Soil Management in the Oxisol Savannas and Ultisol Jungles of Tropical South America". Sanchez summarizes much of the very effective work on the characterization of the soils of these regions and also discusses the field experiments at Brasilia and at Yurimaguas and Pucallpa in Peru. Management of cleared land, with lime, fertilizers, legumes and grass, along lines very similar to those appropriate to the south-eastern USA, has been shown to be successful, whereas it was at the stage of clearance that disastrous erosion could occur; on this evidence, Sanchez' recommendation is that clearance should be carried out gradually by slash-and-burn methods.

As A.J. Smyth says in his foreword, *Characterization of Soils* amply demonstrates the rewards of collaborative research on tropical soils. However readers would be well advised to study it in conjunction with Sanchez' outstanding textbook, *Properties and Management of Soils in the Tropics*, published by Wiley in 1976. □

Harry Vine, who has worked as a soil specialist in Nigeria, was formerly Senior Lecturer in the Department of Geography at the University of Leicester.

The ins and outs of food production

K.L. Blaxter

Biological Efficiency in Agriculture. By C.R.W. Spedding, J.M. Walsingham and A.M. Hoxey. Pp.383. ISBN 0-12-656560-0. (Academic: 1981.) £18.40, \$38.

AGRICULTURE is a very complex undertaking. It involves the use of plants and animals and the deployment of resources of land, labour and capital so to produce food and a range of other products which mankind needs or requires. Using the term "efficient" in the sense "effective", agriculture has obviously been efficient over the centuries for it has resulted in the production of sufficient food to support population growth in the world as a whole and a small but discernible increase in average living standards. Whether agriculture is efficient in the sense of making best use of the resources it deploys is perhaps another question, one to which this study of biological efficiency in agriculture addresses itself.

The major tool employed by the authors in the analysis is that of estimating ratios of the outputs of agricultures per unit time to a number of different inputs, these ratios being termed efficiencies. It is recognized that such single estimates can only apply to specific and highly defined situations, and that the approach can provide little information about the optimization of inputs into agricultural systems. As such, the book is essentially descriptive rather than analytical. It describes crop and animal production systems in many parts of the world in terms of a series of ostensibly simple ratios such as yield per hectare, output per unit of human labour, and yield per unit of fertilizer or water applied and of support energy provided. Annual yields are expressed in terms of dry weight or of energy-yielding constituents edible by man or of protein edible by man. Little emphasis is given to the complexity of the interrelationships between the input terms, that is to response analysis in what must be regarded as a multidimensional system.

A descriptive approach such as that adopted might, nonetheless, have considerable comparative value, so enabling judgements of the relative merits of different crops to be made. The fact that the ratios apply to highly specific contexts of soil, climate, economic and social circumstance, however, makes this impossible. The data assembled in the many tables are interesting as statements about what occurs in farming in different parts of the world. They also have value in predicating a series of other questions in the minds of readers about how to judge the merits of a cropping or stocking system in a particular agricultural setting.

It is wrong to criticize a book because it is about the wrong subject. Nevertheless, a conclusion emerges from this volume which touches on the wider issues, namely

that it is important to decrease the input of support energy into agricultures throughout the world without reducing total food production. Distinctions are made between those uses of support energy in terms of fertilizers and agrochemicals which augment the efficiency with which incident solar radiation is used to produce food and those in which machines simply substitute for manpower. This is a real question which could well have been developed from a closer analysis of the large amount of descriptive material presented. Perhaps such an analysis is already in the minds of the authors. □

Sir Kenneth Blaxter is Director of the Rowett Research Institute, Bucksburn, Aberdeen.

L'animal-machine

J.E.R. Staddon

Quantitative Ethology: The State Space Approach. By David McFarland and Alasdair Houston. Pp.204. ISBN 0-273-08417-8. (Pitman: 1981.) £17.50, \$24.95.

MOST ethologists accept the idea that animal behaviour is to be explained in a purely mechanistic way. What could be more natural, then, but to turn to engineering, the science of mechanism, for the theoretical analysis of behaviour? So it has seemed to David McFarland and his group at Oxford who for the past 15 years or so have pursued the engineering approach with unequalled thoroughness and sophistication. This book is a summary of their work and the work of others in the same tradition.

The work of the Oxford group has gone through two phases. The initial phase was McFarland's application of linear systems analysis to a variety of quantitative motivational problems in which the level of some activity — rate of drinking, or attack, or pressing a lever in a Skinner box, or feather fluffing — was measured as a function of a single independent variable, usually time. This work resulted in *Feedback Mechanisms in Animal Behaviour*, published by Academic Press in 1971.

Simple linear systems analysis represents the state of a system by a single number, but this is obviously an arbitrary restriction. It is much more natural to represent the motivational state of an animal by a point in *N*-dimensional space. This is termed a state-space analysis by engineers. It also corresponds to a mode of representation commonly used by economists. In the motivational case, each axis

might represent the level of some crucial physiological variable; in the economic case, each axis gives the amount consumed of a commodity. In both instances, the benefit to the individual associated with points in the space is represented by a set of equal-benefit contours. *Quantitative Ethology* is largely devoted to descriptions of static and dynamic applications of the state-space approach, the second phase of the Oxford group's programme. The book provides a useful catalogue of methods of analysis and examples of applications. Techniques discussed include the method of Lagrange multipliers (a standard tool for economists), Pontryagin's maximum principle (an approach to dynamic optimization), classical control systems theory, and measurement theory (a discussion of the combination rules describing the collective effect of stimuli and other causal factors on behaviour).

This book differs from McFarland's earlier one in another respect: its emphasis on adaptive function. The formal similarity of the economic and engineering approaches allows the authors to move easily from mechanistic models to optimality models, from an engineering approach where the animal is treated as a control mechanism that seeks to maintain the point representing its internal state within an "acceptable" region, to an economic approach where the organism is assumed to maximize its fitness. In their discussion of decision rules the authors show that they are well aware of the distinction between the limited means animals use to achieve optimal behaviour, and the optimality analyses that define optimal behaviour. Yet the formal similarity between the economic and engineering versions of state-space analysis suggests that further work is necessary to separate the ideal from the means the animal uses to achieve it.

Depending on their definition of "ethology", some will quarrel with the title of the book. If you believe that the distinction between ethology and animal psychology is that the former deals with the ways in which animal species differ from one another, whereas the latter is interested in their similarities, then you will find McFarland and Houston's book closer to psychology than ethology. If you feel that the important questions in ethology are social, or involve ecology or neurophysiology, then the book is not for you. On the other hand, if you consider that our best hope for understanding animal behaviour is to begin with the large and potentially quantifiable processes of motivation and steady-state operant behaviour, then *Quantitative Ethology* presents the clearest exposition to date of the work of the Oxford group on these topics. □

J.E.R. Staddon is a Professor in the Departments of Psychology and Zoology at Duke University.

Fruits of Australian botany

David W. Goodall

Flora of Australia. Vol.1, *Introduction*. Pp.200. ISBN hbk 0-642-06652-3; ISBN pbk 0-642-06653-1. (Australian Government Publishing Service, Canberra: 1981.) Hbk \$A15.50; pbk \$A10.50. *The Vegetation of Australia*. By Noel C.W. Beadle. Pp.690. ISBN 0-521-24195-2. (Cambridge University Press: 1981.) £50, \$120. *Australian Vegetation*. Edited by R.H. Groves. Pp.449. ISBN hbk 0-521-23436-0; ISBN pbk 0-521-29950-0 (Cambridge University Press: 1981.) Hbk £27.50, \$64.50; pbk available in Australia only. *The Biology of Australian Plants*. Edited by J.S. Pate and A.J. McComb. Pp.412. ISBN 0-85564-194-0. (University of Western Australia Press: 1981.) \$A25.

No flora of the whole of Australia has been compiled since George Bentham's *Flora Australiensis*, published from Kew between 1863 and 1878. It is still a standard reference work, despite the enormous amount of taxonomic research undertaken during the intervening century (twice as many species of higher plants are now known in Australia as were recorded by Bentham). State and regional floras have been written but, for important parts of the continent, there has been nothing to replace Bentham — out of date as it is.

Accordingly, it was a major event for Australian botany when, after 20 years' gestation, arrangements were finally completed in 1979 for the publication of a new flora of Australia, of which the first volume is reviewed here. It is planned that *Flora of Australia* will cover the vascular plants in 48 volumes, with an unspecified number of subsequent volumes to deal with the lower plants; attention is at present concentrated on the flowering plants, and practically every angiosperm taxonomist working in Australia will be involved in one or other of the volumes.

The present book is explicitly an introduction to the series. The main elements are an excellent chapter by B.A. Barlow on the origin and evolution of the vascular flora; an outline of the system of classification used, in which A. Kanio compares the recent systems of Takhtajan, Dahlgren, Thorne and Cronquist with more traditional ones; and a key to angiosperm families by H.T. Clifford. The last is based primarily on experience in Queensland, and comments are invited as to how satisfactory it proves on the continental scale. Certainly, this section of the volume is likely to be in frequent use by non-specialists.

A flora is judged, however, not by its introductory material but by the quality of the species descriptions and keys. I have had the opportunity of seeing proofs of another volume in the series, and it is clear that high standards are being set and achieved. Descriptions are clear, including

notes on ecology and geographical distribution (numerous maps are included). The original description and type specimen are identified, and some representative specimens in Australian herbaria are listed.

It is clear that this series will be in constant demand and use wherever serious taxonomic work is in progress. It will be a necessity for all botanical libraries in Australia, and all taxonomic libraries throughout the world.

Two of the other books covered in this review, those of Professor Beadle and Dr Groves, have much in common; both are concerned with vegetation types, have very similar titles and (surprisingly) come from the same publisher. The first, however, is almost entirely one man's work, the second has separate authors for each chapter. Both, but particularly the first, are well illustrated with numerous photographs. Recurring themes in the two volumes are the importance of nutrient status — Australia is a very old landmass, much of which is deficient in various nutrients, particularly phosphorus — and fire as a regular component of the environment to which Australian vegetation has become adapted.

Professor Beadle's compendious volume is the Australian contribution to the series *Vegetationsmonographien der einzelnen Grossräume*. After preliminary chapters on climate, soils and the flora, including its history and origins (cf. the chapter by Barlow in *Flora of Australia*), the various plant communities recognized are described seriatim. The communities are distinguished, named and described on the basis of the dominant species in the over-storey. This makes for numerous units ("associations" or "alliances") and for fragmentation of vegetational continua, in which the same under-storeys may continue under various canopies, and even the dominant species may almost be vicariants, replacing one another with little influence on the rest of the vegetation. Particularly is this so in the vegetation dominated by *Eucalyptus* spp., the description of which occupies one-quarter

The Banksias

Academic Press have recently published the first volume of a three-volume work, *The Banksias*, with paintings by Celia Rosser and text by Alex George.

The project was initiated with the aim of creating a work of both art and taxonomic scholarship. When complete, in 1988, the series will contain illustrations and detailed descriptions of all 72 species of the *Banksia* genus.

Each of the three volumes will be available only in a numbered, limited edition of 720 copies. Price of Vol.1 is £965, \$1,965. The less pecunious may arrange to see the book at the Linnean Society, Piccadilly, London, or at Academic Press offices in London, New York or Sydney.

of the book, in accordance with their importance in the continent. The frequent references to ecotones underline the inadequacy of a treatment based on discrete communities defined by dominants.

The approach adopted implies that the description of a vegetation type largely consists of an account of the geographical distribution and habitat preferences of the dominant species, with a list of the species accompanying it in the under-storey. There are no association tables, and very little on the quantitative make-up of the vegetation. The treatment is mainly descriptive; vegetational dynamics are mentioned in some places, but in general the vegetation is treated as static.

The volume edited by Dr Groves is much less ambitious than that of Professor Beadle, and contains far less floristic detail. Although here, too, there is a tendency to write in terms of dominant-defined communities (as has been the tradition in Australia), some of the writers point out the drawbacks of this approach, and consider alternatives; there is in any case no attempt to treat each community separately. In consequence, the book is far more readable than Professor Beadle's. The different vegetation types are very adequately covered, but in broad groups rather than individually. Although there are inevitable differences in the treatments accorded their subject-matter by the different authors (some chapters are mainly descriptive, others have substantial quantitative components, including estimates of productivity; others again devote attention to analytical and dynamic aspects), the editor has had a good deal of success in achieving uniformity and balance.

These two books will fulfil different functions. Professor Beadle's will serve as a constant reference source on the variety, geographical distribution and floristic make-up of Australian vegetation. Dr Groves's, on the other hand, could well be used as a textbook in advanced courses, particularly where a purely descriptive approach is considered unsatisfactory.

The two volumes share some faults. Neither deals with ruderal or segetal communities, or with the sown and lightly managed grasslands which occupy such large areas in the more densely populated parts of the continent. Little attention is paid to non-vascular components of the vegetation. And, surprisingly, little notice is given to the role of domestic livestock in causing changes and determining the current pattern of Australian vegetation.

The fourth volume reviewed consists of papers presented at a symposium in Perth in 1979. The first four contributions cover plant responses to four types of stress important in Australia — lack of water; low nutrient supply; fire; and salinity. At that point, however, planning of the subject-matter becomes less noticeable, and the topics seem simply to reflect the research interests of participants. This does

not, of course, prevent individual papers being of considerable interest; one may mention, for instance, that by D.E. Gaff on "resurrection plants" — those, such as *Borya nitida*, which can suffer complete desiccation yet resume growth when remoistened (surprisingly, Dr Gaff does not mention lichens and arid-zone mosses, "resurrection plants" *par excellence*); and

that by A.J. McComb *et al.* on seagrasses. The general standard is higher than that of many symposium volumes, and the book should certainly find a deserved place in botanical libraries. □

David W. Goodall is Senior Research Fellow in the CSIRO Division of Land Resources Management, Wembley, Western Australia.

Volcanology: last of the common ground

P. W. Francis

Tephra Studies. Edited by S. Self and R.S.J. Sparks. Pp.481. ISBN 90-277-1327-8. (Reidel: 1981.) DM125, \$54.50.

ICELAND is rather an appropriate venue for international meetings, and especially a NATO-sponsored one, since its position astride the mid-Atlantic ridge makes it mutually accessible for scientists from both North America and Europe. This volume is the outcome of an uncommonly successful NATO Advanced Study Institute held at Laugarvatn, Iceland, in June 1980. Several factors contributed to the success of the meeting: the participants lived and worked together in an isolated hotel, distracted solely by spectacular views of the volcano Hekla on the opposite side of Lake Laugarvatn; the exorbitant cost of liquor in Iceland ensured that most of them were able to arrive at the morning sessions fresh and attentive; and the steamy intimacy of geothermally heated sauna baths meant that ordinary barriers of communication were soon broken down. The constructive discussions that flourished in this environment are reflected in many of the contributions to the volume that resulted.

Tephra is the collective noun used to describe pyroclastic rocks. The book is something of a milestone in volcanology because it is the first to be devoted exclusively to this long neglected but increasingly important subject, which impinges on disciplines as diverse as archaeology, climatology and petrology. It is clearly a volume that libraries and tephra specialists should have on their shelves, but it may be less useful to other classes of reader, precisely because it does include such an extremely broad range of topics — someone interested in marine studies, for example, would derive little benefit from sections on the use of tephra in dating archaeological sites.

Tephra studies are relatively modern. The term was first coined by the Icelandic doyen of the subject, Sigurdur Thorarinsson, to whom the book is most fittingly dedicated. Thorarinsson's introductory chapter is a model of lucid and stimulating technical writing. It is difficult to forget his account of the early Icelandic worker, who, in order to determine the limits of a tephra fall, was accustomed to pick up a pellet of sheep

droppings and take a bite at it. When he could no longer feel grains of fine ash between his teeth, he knew he was outside the zone of tephra falls. While the other 30 or so contributions do not all maintain the same high standard, and are more conventionally turgid, they are generally informative and well presented. Particularly noteworthy are the accounts by Walker of remarkably violent eruptions in New Zealand; the first serious analysis of the effects of heavy ash falls on buildings and other structures by Blong; and the startling suggestion by Steen-McIntyre that archaeological remains at a site in central Mexico, associated with tephra dated by her using the tephra hydration technique, may indicate the arrival of early man in the New World some 150,000 years before the generally accepted date.

Like many other conference proceedings, the book has been produced in camera-ready copy format. The quality of the text is perfectly adequate, but a few of the diagrams have suffered somewhat in the reproduction process from excessive reduction. A few proof-reading errors occur, but much more irritating is the editors' decision not to include the titles of papers in references. Although the space saved by their omission must be significant, so too is the inconvenience caused to readers. Other editors should be urged to reconsider their policy in this respect.

Although the book is an important contribution to the literature of volcanology, it seems unlikely that such a comprehensive study will be repeated soon. There is such a gulf between, for example, the interests and backgrounds of volcanologists studying the mechanisms of emplacement of pyroclastic flows and stratigraphers working on the use of ash layers as marker horizons in the Quaternary, that there is unlikely to be much common ground between them, and future more detailed studies will reduce this still further, until the only common factor will be the pyroclastic origin of the rocks concerned. We must welcome the publication of this book, then, as merely the first of many more specialized ones in the headlong advance of tephra studies. □

P. W. Francis is a Visiting Scientist at the Lunar and Planetary Institute, Houston.